Numerical gradient methods for flux identification in a system of conservation laws

Francois James, Marie Postel

Numerical gradient methods for flux identification in a system of conservation laws

François James [*], Marie Postel [†]

Abstract

The identification of the flux for a system of conservation laws is studied from a numerical point of view, on the specific example of chromatography. Different strategies to compute the exact gradient of the discretized optimization problem are developed and compared. Numerical evidence of the convergence of the method is also given in the scalar and binary case. Finally a ternary mixture with real experimental data is studied and the identified isotherm is compared with chemical engineers results.

Keywords: hyperbolic systems of conservation laws – flux identification – discrete gradient method – chromatography – measure-valued solutions

1 Introduction

The accuracy of any kind of mathematical model relies on the precise knowledge of all the involved parameters, in a wide acceptation. Among these one must think of the initial data, which are only partially known in several applications (meteorology, for instance), and state laws, appearing as nonlinearities in partial differential equations. The inverse problem consists in recovering such data or parameters from experimental observations, in order to improve the current model. Specific problems arise when the governing equations are systems of nonlinear hyperbolic conservation laws, which are involved in numerous examples in physics and chemistry. We focus in this paper on the problem of identifying the flux in a system of conservation laws, motivated by the specific example of chromatography.

The chromatography process is a powerful tool to separate or analyze mixtures. It is widely used in chemical industry (pharmaceutical, perfume and oil industry, ...) to produce relatively high quantities of very pure components. In these conditions, diffusive effects can be neglected, and the behavior can be reasonably modeled by the system of mass balance law (see [15, 14] for more elements concerning models). The process is therefore mainly governed by a nonlinear function of the mixture concentrations, the

[*]Mathématiques, Applications et Physique Mathématique d'Orléans, CNRS UMR 6628, Fédération Denis Poisson, CNRS FR 2964, Université d'Orléans, BP 6759, F-45067 Orléans Cedex 2, FRANCE, francois.james@univ-orleans.fr

[†]Laboratoire Jacques-Louis Lions, CNRS UMR 7598, Université Pierre et Marie Curie, BC 187, F-75252 Paris Cedex 05, FRANCE, postel@ann.jussieu.fr

so-called isotherm function, which appears as the flux of the system. Thermodynamical properties of the isotherm ensure that the resulting system is hyperbolic.
The precise knowledge of the isotherm is crucial, from the theoretical viewpoint of physico-chemical modeling, as well as the more practical preoccupation of accurately governing the experiment to improve separation. Chromatography can be used to identify isotherms, but its application is limited because it requires a rather heavy experimental apparatus. We describe here another approach, which consists in using numerical simulations of the process in order to compare the solutions to experimental outputs.
The problem consists in finding the parameters such that the solution of the model is "as close as possible" to some experimental observation at a given position, which means that a suitable cost function is chosen, typically a least square estimate of the difference between the solution and the observation. We are thus confronted with an optimization problem, for which descent type methods are natural, so that computing the gradient of the cost function is now the problem we focus on. Two strategies can be followed. The first one consists in directly computing the directional derivative of the cost function. It leads to a formula involving the solution to the linearized version of the original system of conservation laws. The other one is a reformulation of the problem in the spirit of control theory, which introduces an adjoint state, solution to a backward system of linear transport equations. A major problem for both formulations is that the coefficients in the linear equations are discontinuous as soon as shocks arise in the nonlinear solution. Thus, a correct formulation for the gradient is presently out of reach.
Therefore we turn to some discrete formulation of the problem, and we follow the strategy which consists in computing the exact gradient of the discretized problem rather than some arbitrary discretization of the continuous formulation. For the above reasons, we do not try to prove any convergence results, but we give numerical evidence that the schemes we obtain behave nicely when refining the discretization. We give also a few elements of comparison between the two strategies.
The paper is organized as follows. In Section 2 we recall the physical context, sketch the main properties of the model, and precisely state the identification problem. Section 3 is devoted to the computation of the gradient, from both the formal and numerical point of view. Next we give numerical results for scalar equations and 2×2 systems (in Section 4) and show an application on a real set of experimental data for a 3×3 system (in Section 5). Finally, some technical computations are gathered in the Annex.

2 Description of the physical problem and model

We recall here the physics and chemistry principles underlying the chromatography process. The separation results from the interaction between two phases in relative movement. The experimental - or industrial - apparatus consists in a column filled with a porous medium in which a neutral solvent circulates at a fixed velocity. A given concentration of the mixture is introduced at the head of the column during a limited time. As the mixture makes its way down the column, part of it is adsorbed at the grain surface and forms what is called the stationary phase. The separation between the different

components of the mixture results from the competition between two phenomena: on the one hand the mobile phase propagates obeying the fluid dynamics laws, on the other hand the balance between the two phases relies on thermodynamical laws from which the notion of diphasic equilibrium is defined.

2.1 Equations for the direct problem

It turns out that the actual experimental or industrial context allows to simplify the hydrodynamical model to an extreme. Indeed, since the length of the column is large compared to its diameter, we can safely neglect all radial effects, thus obtaining a one dimensional model. Next, since the involved concentrations are high, precisely because we wish to observe nonlinear effects, all diffusive phenomena can also be removed. We do not need the energy conservation equation, because the experiment usually takes place at constant temperature (more generally heat transfer effects can be neglected). Finally, the velocity of the mobile phase is assumed to be constant and equal to the vector solvent. This last hypothesis is quite reasonable in the case of liquid incompressible mixtures, and once again this is relevant for lots of chromatography manipulations (HPLC, High Performance Liquid Chromatography). All these hypotheses amount to simplify the physics and leave only the mass conservation to be written.
The thermodynamical model sets to work complicated mechanisms in order to simulate adsorption of the chemical components. Without going into details (see [15, 14]), we merely recall the essential point which consists in supposing the existence and uniqueness of a stable equilibrium state for the thermodynamical system of the two phases.
More precisely, for a mixture of p components, we denote by $\mathbf{c}^1, \mathbf{c}^2 \in \mathbb{R}^p$ with $c_i^j \geq 0$ for $1 \leq i \leq p$ and $j = 1, 2$, the concentrations in phases 1 and 2, with respect to the total volume in the column, of the p chemical components. The equilibrium is modeled by a smooth function $\mathbf{h} : \mathbb{R}^p \to \mathbb{R}^p$, such that $\mathbf{c}^2 = \mathbf{h}(\mathbf{c}^1)$. Furthermore $\mathbf{h}$ has the following properties

$$\mathbf{h}(0) = 0, \tag{1}$$

$$\mathbf{h}'(\mathbf{c}^1) \quad \text{is diagonalizable with eigenvalues} \quad \mu_i(\mathbf{c}^1) > 0. \tag{2}$$

The function $\mathbf{h}$ is called an *isotherm*, which comes from the fact that the local equilibrium is reached at constant temperature. Chemistry literature on isotherms is plentiful (see [15, 14] and bibliography herein). A very classical example of such an isotherm is the *Langmuir isotherm* [19, 20]

$$\mathbf{h}_i(\mathbf{c}) = N^* \frac{K_i \mathbf{c}_i}{1 + \sum_{i=1}^p K_i \mathbf{c}_i}. \tag{3}$$

The model here is completely determined by $p + 1$ parameters, the so-called *Langmuir coefficients* K_i, which are homogeneous to the inverse of a concentration, and the *saturation coefficient* N^*, which corresponds to some limit concentration when the stationary phase is saturated. We shall mainly use this function, or variants of it, for our numerical simulations throughout the paper.

In order to write the mass conservation equation we consider the two phases: the mobile one - phase 1 - moves downward with a speed $u > 0$; the adsorbed phase - phase 2 - has null speed $v = 0$. Therefore we have

$$\partial_t(\mathbf{c}^1 + \mathbf{c}^2) + \partial_x(u\mathbf{c}^1) = 0. \tag{4}$$

In this experimental setting, we can assume that the equilibrium between the two phases is instantaneous (quasi-static process). At all times and everywhere in the column we can write the closing relation between the concentrations in the two phases, using the isotherm. Equation (4) can then be rewritten using only the mobile phase concentrations $\mathbf{c} \equiv \mathbf{c}^1$

$$\begin{cases} \partial_x \mathbf{c} + \partial_t \mathbf{F}(\mathbf{c}) = 0, \quad t \in [0, T], \quad x \in [0, L], \\ \mathbf{c}(0, t) = \mathbf{c}_{inj}(t), \\ \mathbf{c}(x, 0) = 0, \end{cases} \tag{5}$$

where the function $\mathbf{F}$ is given by

$$\mathbf{F}(\mathbf{c}) = \frac{1}{u}\left(\mathbf{c} + \frac{1-\varepsilon}{\varepsilon}\mathbf{h}(\mathbf{c})\right), \tag{6}$$

and $0 < \varepsilon < 1$ is the void fraction of the column.
Because of the properties of the isotherm (2), the system (5) is hyperbolic. Notice that in (5), the time and space variables exchange their usual role: the evolution variable is x here. This trick avoids the inversion of the function $\mathbf{F}$ during simulations, and is made possible because the eigenvalues λ_i of $\mathbf{F}'$ turn to be positive (this is a direct computation using the positivity of the μ_i-s in (2)). This ensures that characteristics always enter the orthant $\{t > 0, x > 0\}$. To be a little bit more specific, notice that one can check that $0 < \lambda_i < u$ for all i, so that the "concentration waves" propagate with a smaller velocity than the inert tracer: we actually model a retention phenomenon.
The injection of the mixture takes place at the head of the column during a limited time: this boundary condition is given along with experimental data and has more or less the shape of a notch. As an example Figure 1 shows the concentrations measured at the output of the column as a function of time: these curves are called chromatograms. The mixture under consideration has three different components (BA, PE, MBA) injected in proportions $1 : 3 : 1$ with the time injection profile denoted by 'inj (BA)' in the figure.
For each component the experimental concentrations are displayed with symbols and the concentrations computed with the model (5) and a special isotherm function provided by the chemical engineers are displayed with lines (see (23), Table 5 and [24]). In this experiment the component BA (□) has been clearly separated from the other two (△ and ○) with a peak in concentration reaching the bottom of the column half a minute ahead of the other two. Another remarkable feature is the strong hyperbolic behavior of the chromatograms with clearly identifiable shocks. This is due to the high level of concentrations which makes this dataset a very interesting benchmark, which will be further studied in Section 5.

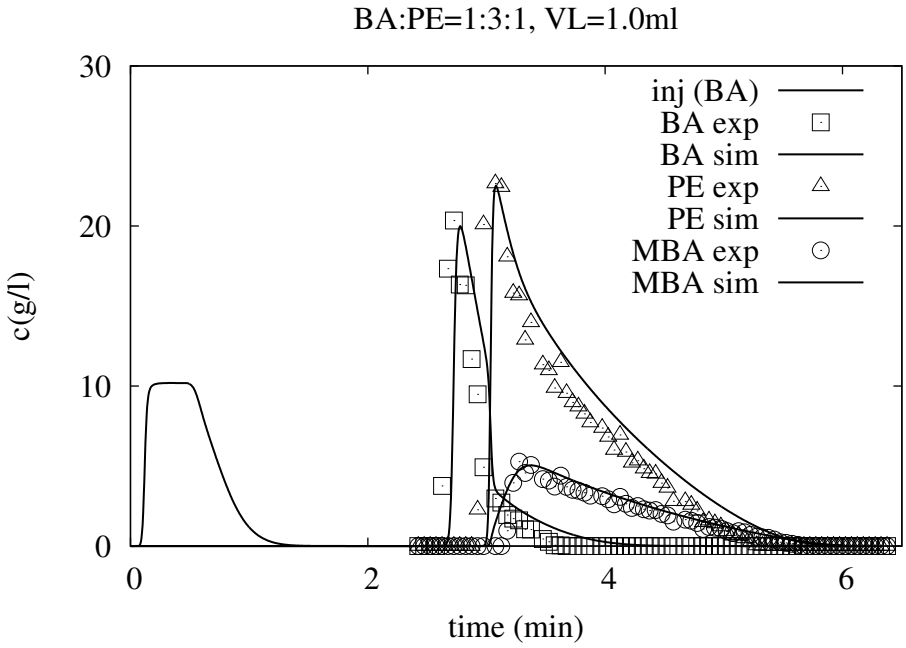

Figure 1: Concentrations in BA, PE and MBA components at the column output.

2.2 Identification

A problem of major practical interest for chemical engineers consists in identifying the physical parameters of the isotherm. Chromatography can be used as a tool to obtain experimental values of pairs $(\mathbf{c}, \mathbf{h}(\mathbf{c}))$. Several methods can be used for this, for a synthetic review, see [18], and for a more complete discussion [15, 14]. For competitive isotherms, that is $\mathbf{c} \in \mathbb{R}^p$ with $p > 1$, the so-called *Frontal Analysis* (FA) methods are the only available. However, they are very slow and require significant amounts of usually expensive pure chemicals. Furthermore, to obtain relevant information for a function from $\mathbb{R}^p$ to $\mathbb{R}^p$, one needs a significant number of such pairs $(\mathbf{c}, \mathbf{h}(\mathbf{c}))$, and each of them results from a heavy experiment. For instance, the isotherm in Section 5 below has been identified by a FA method from a set of 30 experimental points. The PDE model is not used here, the parameters are obtained by direct fitting on the experimental measurements of isotherms.

The alternative approach we use here consists in using as observation the concentration profiles with respect to time at the output of the column (chromatograms). The underlying principle is quite simple: one measures the concentrations in the mixture at the exit of the column. The fit between these experimental data and the model is measured through a cost function $J(\mathbf{F})$, which we choose here as the classical least squares

observation:

$$J(\mathbf{F}) \quad = \quad \frac{1}{2}\int_0^T \|\mathbf{c}_\mathbf{F}(L,t) - \mathbf{c}_{obs}(t)\|^2\, dt \tag{7}$$

where $\mathbf{c}_\mathbf{F}$ denotes the solution to (5), and $\|.\|$ is some norm on $\mathbb{R}^p$. This formulation presumes that we have a complete knowledge of each component of the mixture, as in Figure 1. If the experimental apparatus does not provide such a measure, the observation is then the total concentration of the mixture, say $\mu_{obs}(t)$, that is the sum of the different components. In that case, the function J can be chosen as

$$J(\mathbf{F}) \quad = \quad \frac{1}{2}\int_0^T |\sum_{i=1}^{p} \mathbf{c}_{\mathbf{F}i}(L,t) - \mu_{obs}(t)|^2\, dt.$$

For instance, had the dataset in Figure 1 be recorded as the sum of all three components concentrations, the separation of the last two components would have been hardly recognizable.

However, thanks to a higher experimental price, one can obtain such experimental datasets where the concentrations could be measured separately, which of course enhances a lot the sensitivity of the cost function (7) with respect to the isotherm parameters. We will always use this type of datasets, artificially generated by numerical simulation in Section 4, or experimentally obtained, in Section 5.

At this point, we emphasize the fact that, for practical applications, one cannot expect to identify the isotherm directly as a function from $\mathbb{R}^p$ to $\mathbb{R}^p$. Indeed, the problem is severely ill-posed as soon as shocks are present in the observation, and no uniqueness is ensured. Instead we prefer to identify the parameters of some given analytical model, which enjoy a convenient physical interpretation, and sometimes can be roughly predicted from experimental data. For instance, consider the Langmuir isotherm (3). It is completely determined by the knowledge of the coefficients K_i and N^*. On the one hand, N^* is a concentration at saturation for the porous medium, on the other hand, the products N^*K_i can be more or less accurately estimated from the data by picking on the chromatograms the second crossing of each component of the concentration with the time axis. These times correspond to the end of the rarefaction wave and are related with the gradient of the flux function as

$$T_i = \frac{L}{u}\left(1 + \frac{1-\varepsilon}{\varepsilon}\frac{\partial \mathbf{h}}{\partial \mathbf{c}_i}(0)\right).$$

If we denote by $\alpha = (\alpha_1, \ldots, \alpha_q)$ the parameters of the isotherm to be identified, the functional $J(\mathbf{F})$ becomes a function $\tilde{J}(\alpha)$ from $\mathbb{R}^q$ to $\mathbb{R}_+$, which has to be minimized over some subset of $\mathbb{R}^q$. For the Langmuir isotherm, we have $q = p + 1$, but for more realistic (and complex) models q can be quite larger than p.

Notice finally that the minimization problem as it is stated definitely does not fall into classical settings. The function J has no reason to be convex, actually numerical results show that there are local minima. Differentiability is an open problem (see below Section 3). Thus one can think to use nonlocal optimization methods to address this problem. There were indeed some attempts for chromatography, using genetic algorithms (see [11, 9]). These methods present the advantage of requiring few a priori

knowledge on **F** (in particular they do not use any kind of derivative), but their main drawback is that they require a large number of direct simulations of the numerical model.

2.3 Discrete formulation

We now turn to the discrete version of the optimization problem, since the experimental data will be provided at a discrete sampling rate. Furthermore the direct model cannot be solved in its continuous version (5). We choose to obtain an approximate solution by discretizing and solving it numerically using a standard finite volume method, well adapted to this type of hyperbolic system.
We define a uniform grid in time and space

$$\begin{aligned} x_0 = 0 < x_1 = \Delta x < \ldots < x_k = k\Delta x < \ldots < x_K = K\Delta x = L, \\ t_0 = 0 < t_1 = \Delta t < \ldots < t_n = n\Delta t < \ldots < t_N = N\Delta t = T, \end{aligned}$$

where we will compute the solution using a Godunov scheme:

$$\mathbf{c}_{k+1}^n = \mathbf{c}_k^n - \lambda \left(\mathbf{F}(\mathbf{c}_k^n) - \mathbf{F}(\mathbf{c}_k^{n-1}) \right). \tag{8}$$

Here $\mathbf{c}_k^n$ is an approximation of the mean value of the solution $\mathbf{c}$ at points $x_k = k\Delta x$ and times $t_n = n\Delta t$. In particular $\mathbf{c}_0^n$ is an approximation of the injection condition in $x_0 = 0$. Similarly, the initial condition $\mathbf{c}(x, t = 0) = 0$ is discretized as $\mathbf{c}_k^0 = c_{initial}(x_k)$ for $k = 0, ...K$.
There is no difficulty with this scheme which is known to be of order one in time and space (see [12]) given the nice monotonicity properties of the flux function **F**: since all eigenvalues of $\mathbf{F}'$ are positive, we are left with a simple upwind scheme. Remark however that contrarily to standard use, it is here the space variable, that is the abscissa along the column, which plays the role of evolution variable in the numerical scheme.
The cost function to be minimized can be obtained by discretizing (7) on this spatial grid:

$$\tilde{J}(\mathbf{F}) = \frac{1}{2}\Delta t \sum_{n=1}^{N} \|\mathbf{c}_K^n - \mathbf{c}_{exp}(t_n)\|^2. \tag{9}$$

One should note however that the definition of the discrete cost function is not as straightforward as it seems: first of all, the optimization relies on the assumption that the system (5) correctly models the experiment, hence there exists a function **F**, completely determined by a set of parameters α, for which the experimental data is solution of it – up to measurements precision. Therefore, in the ideal case where the measurements can be obtained with a sampling rate as small as possible and where the exact solution $\mathbf{c}(x, t; \alpha)$ of (5) can be computed for any set of parameters α, the cost function is

$$\tilde{J}(\alpha) = \frac{1}{2}\Delta t \sum_{n=1}^{N} \|\mathbf{c}(x_K, t_n; \alpha) - \mathbf{c}_{exp}(t_n)\|^2. \tag{10}$$

One should therefore question the convergence of the minimization algorithm when the sampling rate Δt goes to zero. Do the parameters α minimizing (10) tend, to the limit and in a sense to be defined, towards the parameters α which minimize (9)?
This is a very difficult problem, and in any case hardly useful from the practical point of view since experimental data are far from perfect. Furthermore they are sampled on a time grid most often very coarse with respect to the sampling rate necessary in order to get a good numerical approximation of (5) using (8).
Therefore a good way of measuring the fit of the experiment and the model is to fix the discretization of the cost function J equal to the sampling rate of the experimental data, but to compute the model approximations of $\mathbf{c}(x_K, t_n; \alpha)$ with the numerical scheme, with a discretization fine enough to ensure good convergence. This implies a resampling on a coarser grid of numerical solution and of their gradient which is not yet implemented in practice. The state of the art consists in interpolating the experimental data on the fine numerical grid using linear interpolation.
Given the high theoretical difficulty of the convergence of the optimization process, we will address in this study two intermediate problems which are interesting to understand the minimization convergence and that we pose directly at the discretized level. Both will be illustrated by numerical simulations.
First we define a discrete cost function

$$\hat{J}(\alpha) = \frac{1}{2}\Delta t \sum_{n=1}^{N} \|\mathbf{c}_K^n - \hat{\mathbf{c}}_K^n\|^2, \tag{11}$$

where $\hat{\mathbf{c}}_K^n$ is the discrete solution computed with scheme (8) with α set equal to the target value which we denote by $\hat{\alpha}$. This functional has a global minimum in $\hat{\alpha}$ and we will numerically illustrate that our minimization method correctly handles this simple case, and is numerically stable when we let Δt go to 0.
In a second problem, we define the cost function as

$$\tilde{J}_\delta(\alpha) = \frac{1}{2}\Delta t \sum_{n=1}^{N} \|\mathbf{c}_K^n - \mathbf{c}_\delta(t_n)\|^2, \tag{12}$$

where the "experimental data" $\mathbf{c}_\delta$ is the numerical solution computed with the scheme (8) with α set equal to the target values $\hat{\alpha}$ and a very fine discretization in time and space δ, so that it correctly mimics the exact continuous solution of (5). The values $\mathbf{c}_\delta(t_n)$ are obtained by resampling this solution on the coarser grid of sampling Δt. This time, the global minimum of $\tilde{J}_\delta$ is not trivially obtained in $\hat{\alpha}$, except for the limit case where $\Delta t = \delta$, where we are back to the previous cost function (11) where $\tilde{J}_\delta(\hat{\alpha}) = 0$. Here again we will numerically illustrate the convergence of the minimization algorithm as Δt goes to δ.

3 Gradient computation

As soon as there is more than one parameter to identify, it is necessary to be able to estimate the gradient of the cost function with respect to these parameters to ensure

a good behavior of the descent methods. Finite difference estimation of the partial derivatives can be used but they introduce additional sampling rates - in the parameters directions - which have to be calibrated.
The gradient of the functional J with respect to $\mathbf{F}$ (or with respects to parameters α) is hard to study directly starting from the continuous formulation (7)-(5). However, in order to understand the kind of objects we have to deal with, we first perform some formal computations at the continuous level, and present two possible strategies to get a formulation of the gradient. As we shall see, the equivalence of the two formulations is an open problem, and this suggests that the functional is not differentiable in general. Notice by the way that, in the scalar case, it is in some sense Lipschitz continuous. Indeed, several authors, see for instance Lucier [21], Bouchut and Perthame [7] for conservation laws, Evje & al. [10] for degenerate parabolic equations, provide L^1 dependence results of the entropy solution with respect to the flux. Application of Hölder's inequality gives then the result for J.

3.1 Continuous equations

First, we compute some kind of directional derivative of $J(\mathbf{F})$: let $\delta\mathbf{F}$ be some admissible direction, at this level one can think of any smooth function. For $\lambda > 0$, we denote by $\mathbf{c}$ (resp. $\mathbf{c}_\lambda$) the solution to (5) corresponding to the flux $\mathbf{F}$ (resp. $\mathbf{F} + \lambda\delta\mathbf{F}$), so that

$$\frac{J(\mathbf{F}+\lambda\delta\mathbf{F}) - J(\mathbf{F})}{\lambda} = \int_0^T \left\langle \frac{\mathbf{c}_\lambda(L,t) - \mathbf{c}(L,t)}{\lambda}, \frac{\mathbf{c}_\lambda(L,t) + \mathbf{c}(L,t)}{2} - \mathbf{c}_{obs}(t) \right\rangle dt. \qquad (13)$$

Passing to the limit $\lambda \to 0$ in this relation is an open problem in general. Indeed on the one hand, we can expect that

$$\frac{\mathbf{c}_\lambda(L,.) + \mathbf{c}(L,.)}{2} - \mathbf{c}_{obs} \to \mathbf{c}(L,.) - \mathbf{c}_{obs} \qquad \text{in} \quad L^1_{loc}(\mathbb{R}),$$

and this actually holds for scalar conservation laws, see [21, 7]. On the other hand, if $(\mathbf{c}_\lambda(L,t) - \mathbf{c}(L,t))/\lambda$ has a limit, say $\delta\mathbf{c}$, it turns out that $\delta\mathbf{c}$ has to solve the linearized equation

$$\begin{cases} \partial_x \delta\mathbf{c} + \partial_t\Big(\partial_{\mathbf{c}}\mathbf{F}\delta\mathbf{c}\Big) + \partial_t\delta\mathbf{F}(\mathbf{c}) = 0, \quad t \in]0,T[, \quad x \in]0,L[, \\ \delta\mathbf{c}(0,t) = 0, \qquad \delta\mathbf{c}(x,0) = 0, \end{cases} \qquad (14)$$

where $\partial_{\mathbf{c}}\mathbf{F}$ is the matrix of the partial derivatives of $\mathbf{F}$ with respect to the concentrations $\mathbf{c}$

$$(\partial_{\mathbf{c}}\mathbf{F})_{i,j} \quad = \quad \left(\frac{\partial \mathbf{F}_i}{\partial \mathbf{c}_j}\right). \qquad (15)$$

The trouble here is that the solutions to (14) take values in the space of measures in t, so that the only convergence one can hope is too weak to deal with the product in (13). Several notions of solutions have been developed in this context, in the scalar case. See, in the context of multidimensional transport equations, DiPerna-Lions [8], Ambrosio [1] and Bouchut-James-Mancini [6]. For conservation equations, which are

involved here, see Bouchut-James [4] in the one-dimensional case, Poupaud-Rascle [23] for multidimensional results. Notice that DiPerna-Lions and Ambrosio work with renormalized solutions, and assume (roughly speaking) the divergence of the velocity field to be integrable. This is not the case here, because the velocity involves the derivative of $\mathbf{c}$, which can be a nonpositive measure if shocks are present. One can find a justification of weak convergence property in the scalar case, in the setting of differentiation with respect to initial data, see [5]: it makes use of the duality solutions developed in [4].
However, one can say that, at least formally, derivation in the direction $\delta\mathbf{F}$ leads to

$$J'(\mathbf{F})\,\delta\mathbf{F} = \int_0^T \langle \mathbf{c}(L,t) - \mathbf{c}_{obs}(t), \delta\mathbf{c}(L,dt)\rangle\,,$$

where $\delta\mathbf{c}$ is some solution to the system (14). The first strategy to compute the gradient of J consists therefore in finding some numerical evaluation of (16) and (14). Discretizing (14) is not straightforward, we refer to [13] for some results in the scalar case.
Notice that, if the function $\mathbf{F}$ depends explicitly on a number q of parameters $(\alpha_1,\dots,\alpha_q)$, then, instead of computing the derivative of J in some direction $\delta\mathbf{F}$ which is not clearly defined, we have to compute the gradient of J with respect to the α_j-s. Therefore we consider $p\times q$ admissible directions, given by

$$\begin{aligned} \delta\mathbf{F} &= \partial_\alpha \mathbf{F} \quad \text{(in matrix form)},\\ (\delta\mathbf{F})_{i,j} &= \left(\frac{\partial \mathbf{F}_i}{\partial \alpha_j}\right), \quad i=1,\dots,p, \quad j=1,\dots,q, \end{aligned} \tag{16}$$

and the system (14) has to be interpreted now as a matrix-valued equation, where the components of the unknown $\delta\mathbf{c}$ are

$$(\delta\mathbf{c})_{ij} = (\partial_\alpha \mathbf{c})_{ij} = \frac{\partial \mathbf{c}_i}{\partial \alpha_j}, \quad i=1,\dots,p, \quad j=1,\dots,q.$$

Explicit expressions for $\delta\mathbf{F}$ and $\partial_{\mathbf{c}}\mathbf{F}$ are given in Annex III below, for different isotherms.
Another possible expression for $J'(\mathbf{F})$ is obtained by reinterpreting the minimization problem

$$\min_{\mathbf{F}} \frac{1}{2}\int_0^T \|\mathbf{c}_{\mathbf{F}}(L,t) - \mathbf{c}_{obs}(t)\|^2\,dt$$

as a constrained minimization problem:

$$\min\left\{\tilde{J}(\mathbf{v}) = \frac{1}{2}\int_0^T \|\mathbf{v}(t) - \mathbf{c}_{obs}(t)\|^2\,dt, \quad \mathbf{v}(t) = \mathbf{c}_{\mathbf{F}}(L,t) \quad \text{solution to (5)}\right\}. \tag{17}$$

This formulation is rather classical in control theory and parameter identification, and previous results on chromatography are based on it (see [25, 16, 18]). For the sake of completeness, we propose a detailed computation in Annex I below, and merely recall the results here.

The Lagrange multiplier corresponding to the constraint turns out to be a function $\mathbf{p}(x,t)$, solution to the backward linear transport equation

$$\begin{cases} \partial_x \mathbf{p} + \partial_{\mathbf{c}} \mathbf{F}(\mathbf{c})^T \partial_t \mathbf{p} = 0, \quad t \in]0,T[, \quad x \in]0,L[, \\ \mathbf{p}(L,t) = \mathbf{c}(L,t) - \mathbf{c}_{obs}(t), \\ \mathbf{p}(x,T) = 0. \end{cases} \tag{18}$$

Using the fact $\tilde{J}(\mathbf{c}_{\mathbf{F}}(L,t)) = J(\mathbf{F})$, and the Lagrangian corresponding to the constraint, we get another formula for the gradient of $J(\mathbf{F})$:

$$J'(\mathbf{F})\delta\mathbf{F} = \int_0^L \delta\mathbf{F}(\mathbf{c}_{ini}(x))\mathbf{p}(x,0)\,dx + \int_0^L \int_0^T \delta\mathbf{F}(\mathbf{c})\partial_t \mathbf{p}\,dt\,dx \tag{19}$$

for any $\mathbf{p}$ solution to the adjoint equation (18).
The equivalence between (19) and (16) is justified only for smooth solutions, or under specific assumptions in the scalar case, see [17]. See also Bardos and Pironneau [2] in the case of differentiation with respect to the initial condition. A specific problem when discontinuities occur in the solution $\mathbf{c}$ to (5) is that uniqueness is not ensured for the backward problem (18). Therefore stability problems can arise when discretizing the equations. Also, as mentioned in [2], a crucial point to prove equivalence is a convenient definition of the nonconservative product $\partial_{\mathbf{c}}\mathbf{F} \times \delta\mathbf{c}$, when $\delta\mathbf{c}$ is measure-valued. This may indicate that the functional J is not differentiable in general. The nonuniqueness of the adjoint equation might be related to the notion of subdifferential. For a partial result in this direction, see [17], where the convergence of a viscous regularized problem is studied. This remains a very interesting and difficult open question.

3.2 Discrete version

We propose now a strategy to obtain a numerical approximation for both formulations of the gradient, and we also give some elements of comparison between them. The key strategy here is to compute the exact gradient of the discretized problem, instead of applying arbitrary schemes to the above systems of PDE-s. Actually, we start from a given scheme for the direct problem (5), and then mimic the derivation of the continuous formulæ for the gradient. This actually provides numerical schemes both for (14) and (18), and it turns out that the numerical behavior is good. A reason for this could be that such discretizations implicitly define a nonconservative product which is in some sense consistent with the equations. The detailed understanding of this phenomenon, as well as the convergence of the discrete objects remain open problems up to now. Concerning the adjoint formulation, a few steps in this direction can be found in [17]: we have convergence of the sequence of discrete gradients, in the scalar case, and provided the discontinuities of $\mathbf{c}_{obs}$ are exactly observed. The situation is worse for the direct formulation, since we only have some consistency and convergence results for numerical schemes for (14), once again in the scalar case, see [13].
We proceed now to the derivation of the schemes, following the same strategy as in the continuous case. Therefore we compute some exact gradient of the discretized problem. We have the following lemma, which is proved in Annex II.

Lemma 1 *The gradient of $\tilde{J}$ is given by*

$$\tilde{J}'\delta\mathbf{F} \equiv \lim_{\alpha\to 0} \frac{\tilde{J}(\mathbf{F}+\alpha\delta\mathbf{F}) - \tilde{J}(\mathbf{F})}{\alpha} = \Delta t \sum_{n=0}^{N} \left(\mathbf{c}_K^n - \mathbf{c}_{exp}(t_n)\right)^T \delta\mathbf{c}_K^n, \tag{20}$$

where $\delta\mathbf{c}_k^n$ is computed with the scheme

$$\delta\mathbf{c}_{k+1}^n = \delta\mathbf{c}_k^n - \lambda\left(\left(\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^n)\right)\delta\mathbf{c}_k^n - \left(\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^{n-1})\right)\delta\mathbf{c}_k^{n-1}\right) - \lambda\left(\delta\mathbf{F}(\mathbf{c}_k^n) - \delta\mathbf{F}(\mathbf{c}_k^{n-1})\right), \tag{21}$$

with $\delta\mathbf{c}_0^n = 0$ and $\delta\mathbf{c}_k^0 = 0$ as initial and boundary conditions.

By construction, the values obtained for the gradient using this method are rigorously equal to the values obtained using the constrained formulation used until now for this problem (see [25, 16, 18]).

$$\tilde{J}'(\mathbf{F})\delta\mathbf{F} \quad = \quad -\lambda\sum_{n=1}^{N-1}\sum_{k=0}^{K-1}\delta\mathbf{F}(\mathbf{c}_k^n)\left(\mathbf{a}_{k+1}^n - \mathbf{a}_{k+1}^{n+1}\right) + \lambda\sum_{k=0}^{K-1}\delta\mathbf{F}(\mathbf{c}_k^0)\mathbf{a}_{k+1}^0 - \lambda\sum_{k=0}^{K-1}\delta\mathbf{F}(\mathbf{c}_k^N)\mathbf{a}_{k+1}^{N-1},$$

where $\mathbf{a} = (\mathbf{a}_k^n)_{k,n}$ is given by the so-called adjoint scheme of (8) (see Annex I)

$$\begin{cases} \mathbf{a}_k^n = \mathbf{a}_{k+1}^n - \lambda\left(\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^{n+1})\right)^T\left(\mathbf{a}_{k+1}^n - \mathbf{a}_{k+1}^{n+1}\right), \quad 0 \le k < K, 0 \le n \le< N, \\ \mathbf{a}_K^n = \Delta t\left(\mathbf{c}_K^{n+1} - \mathbf{c}_{exp}(t_{n+1})\right), \quad 0 < n < N, \\ \mathbf{a}_k^N = 0, \quad 0 \le k \le K. \end{cases} \tag{22}$$

Although it is difficult to justify theoretically the computation of the gradient using the measure equation (14) when the discretization goes to 0, its numerical behavior is very stable and can be interpreted in term of approximation of delta function as it will be seen in the numerical simulations in following section.
Since they are equivalent in term of accuracy it is interesting to compare the two schemes in term of numerical complexity. In the "direct derivation" method, the unknown $\delta\mathbf{c}$ in the problem (21) is a $p \times q$ matrix which can be computed along with the direct problem unknown $\mathbf{c}$ which is a p vector. On the other hand the unknown $\mathbf{a}$ in the adjoint method is only a p vector - instead of a $p \times q$ matrix - but its computation requires to store the direct problem solution $\mathbf{c}$ for the N times and K abscissa, since the scheme (22) has final boundary conditions in time and space and its coefficients $\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^n)$ depend on the direct scheme solution. The memory requirement is then $N \times p \times q$ for the *direct derivation* method and $N \times K \times p$ for the *adjoint* one. For one evaluation of the gradient, both methods require $NK(p+q)$ calls to isotherm dependent functions $\delta\mathbf{F}$ and $\partial_{\mathbf{c}}\mathbf{F}$. In terms of elementary operations they also require of the order of $N\times K\times p\times(q+p)$ multiplications in the *adjoint* case against $N\times K\times q\times p^2 + N\times q\times p$ in the *direct derivation* case. The leading order term is in both cases an $O(p \times N \times K)$ with a factor $q + p$ in the *adjoint* case and a factor $q \times p$ in the *direct derivation* case. We estimate that for reasonably small values of the number of parameters q, the huge memory requirement of *adjoint* method makes it prohibitive compared to the *direct derivation*. Another argument in favor of the direct computation is the possibility of computing both the solution $\mathbf{c}$ and its derivatives $\delta\mathbf{c}$ on the same adaptive grid, while in

N^*	100.	U	3
K'	5	Inj	25
FlowRate	1	Porosity	0.59
CFL	0.8	# time steps	4000
Diameter	0.39	Length	15.

Table 1: Parameters of simulation in the scalar case

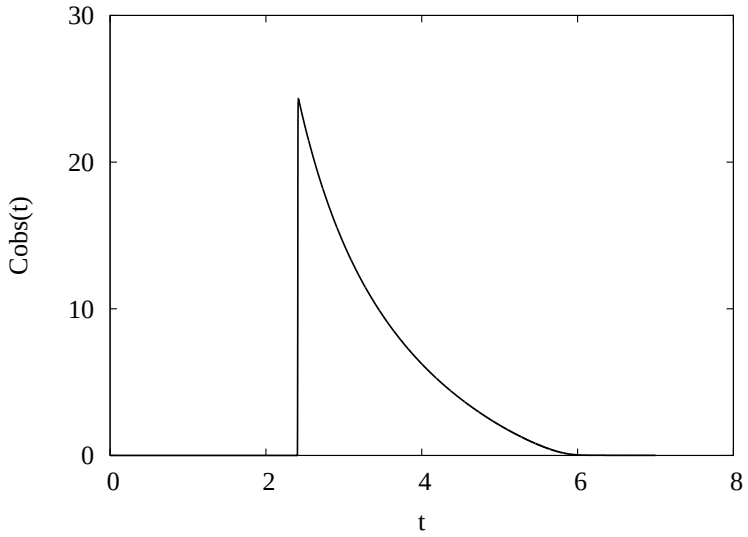

Figure 2: Chromatogram simulated with Table 1

the adjoint method the solutions of the direct and adjoint problems can have discontinuities in different places. The time varying adaptive grid devised to compute the direct problem cannot be easily used to compute the adjoint solution.
On the other hand if the goal of the identification problem was the injection profile, instead of the isotherm parameters, the number of unknowns would become large enough to make the adjoint method more efficient than the direct derivation one (see [2, 22, 3]).

4 Numerical verification of the convergence

4.1 Scalar case

For this set of experiments we have simulated a chromatogram using a Langmuir isotherm with the parameters in Table 1
The chromatogram is displayed on Figure 2. Using this as experimental data, we perform several identification of the isotherm parameters using different space and time discretizations, from 200 to 4000 time steps. When we use the same discretization as the one used to generate the dataset, we should and actually do recover the exact

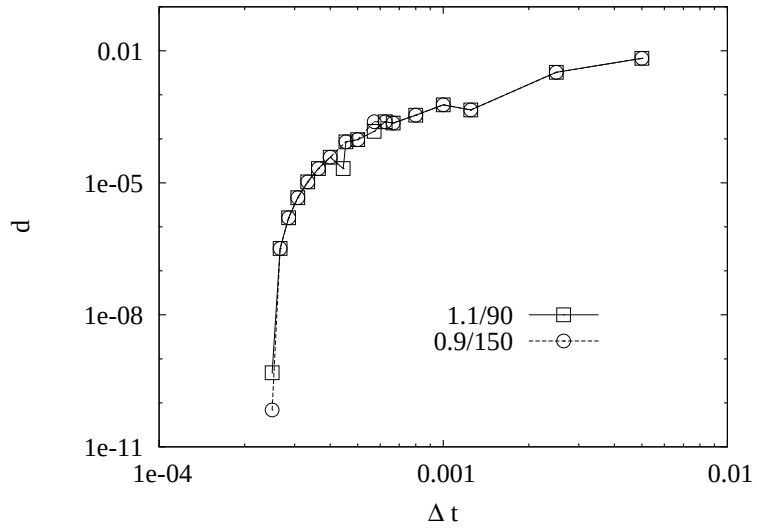

Figure 3: Distance of the optimization result to the target as a function of discretization parameter Δt.

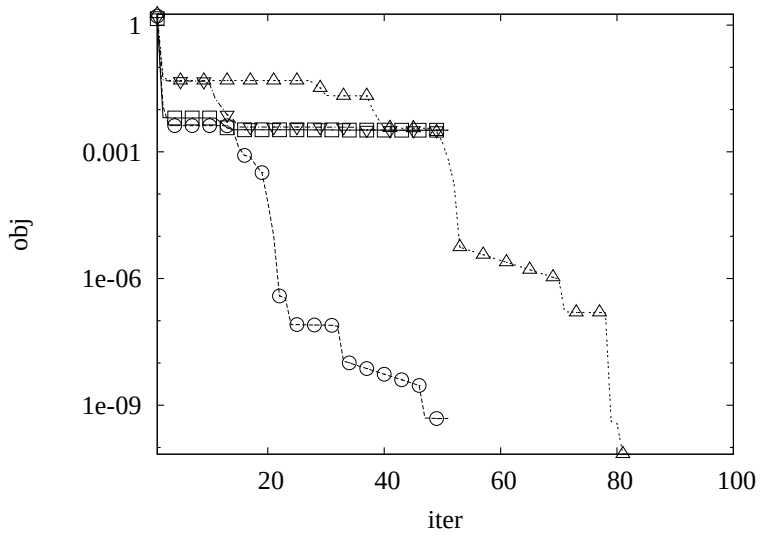

Figure 4: Convergence paths for $Nt = 400, 4000$ and the two initial guesses.

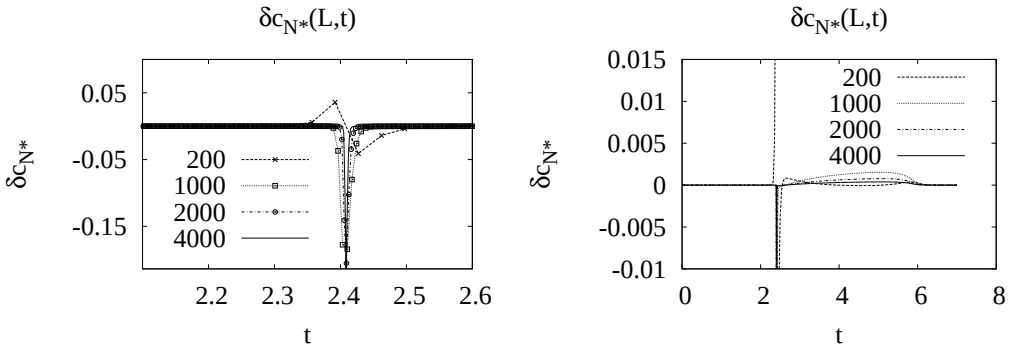

Figure 5: First component of the measure solution $\delta c_1(L,t)$ for $Nt = 200, 1000, 2000$ and 4000 times steps, convergence isotherm used in each case.

parameters of the isotherm. For coarser discretization the gradient identification converges to a target slightly away from the true parameters but which converges towards it with the discretization. This behavior is displayed on Figure 3 which represents in logarithmic scale the distance between the result of the identification and the true target in the parameters space (K' and N^*) as a function of the time step.
The two curves correspond to identifications starting from different initial guesses ($\square$ for $K' = 1.1$, $N^* = 90$) and ($\bigcirc$ for $K' = 0.9$, $N^* = 150$). This figure shows that the descent algorithm converges towards a minimum independent of the starting point, and which gets closer to the theoretical target as the discretization is refined. Figure 4 displays the evolution of the objective function with the iterations for two different discretizations and the two initial guesses. When starting from ($K' = 1.1$, $N^* = 90$), $\square$ indicate the behavior of the minimization done using 400 time steps and $\bigcirc$ correspond to the 4000 time steps computation. The minimization starting from the other initial guess ($K' = 0.9$, $N^* = 150$) is displayed with $\triangledown$ for the 400 time steps computation and $\triangle$ for the 4000 one. The two coarse grid computations ($\square$ and $\triangledown$) converge to roughly the same level of 0.003 for the objective function, while the fine grid computations ($\bigcirc$ and $\triangle$) reach very small values below 10^{-9}. For both discretizations the first initial guess ($K' = 1.1$, $N^* = 90$ indicated by $\square$ and $\bigcirc$) leads more rapidly to the convergence state than the other initial guess indicated by the triangles.
Using the results of this convergence study, we can also perform a closer analysis of the numerical behavior of the measure equation (21). For four different discretizations ($Nt = 200, 1000, 2000$ and 4000) we compute the solution to (21) using the best isotherms parameters for the corresponding discretization, as recorded in Table 4.1. It is a vector of two components corresponding to the derivatives of the concentration with respect to the two parameters of the isotherms K' and N^*. We display on Figure 5 the first coordinate at the output of the column $\delta \mathbf{c}_{N^*}(L,t)$, on Figure 6 the second one $\delta \mathbf{c}_{K'}(L,t)$ and on Figure 7 the difference between the simulated and experimental chromatograms $\mathbf{c}(L,t) - \mathbf{c}_{obs}(t)$. The left hand side graphs display a zoom in the time range of interest. When the convergence isotherm is used the derivatives converges theoretically towards a delta function located at the shock position ($t = 2.41$) and this

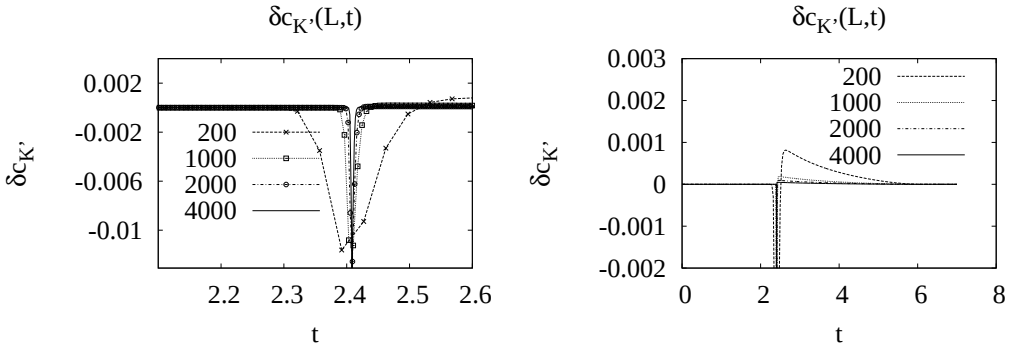

Figure 6: Second component of the measure solution $\delta c_1(L,t)$ for $Nt = 200, 1000, 2000$ and 4000 times steps, convergence isotherm used in each case.

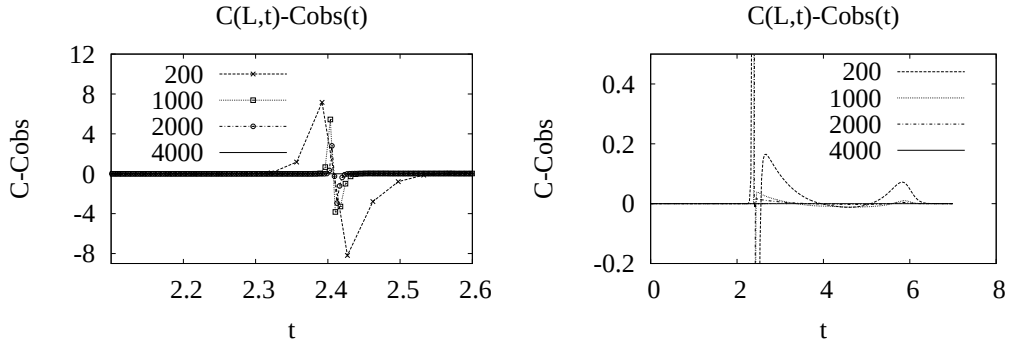

Figure 7: Difference between the simulated and experimental chromatograms for $Nt = 200, 1000, 2000$ and 4000 times steps, convergence isotherm used in each case.

Nt	K'	N^*
200	4.91223	97.5226
1000	4.9804	99.5889
2000	4.99402	99.8801
4000	5.	100.00

Table 2: Best parameters for different discretization in the scalar case

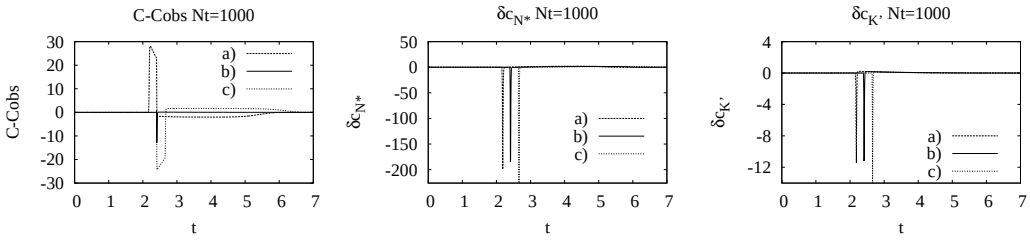

Figure 8: $c(L,t) - C_{obs}(t)$ and $\delta c(L,t)$ for $Nt = 1000$ time steps, and three parameter sets (N^*, K') a) $(90,4)$ b) $(100,5)$ and c) $(110,6)$.

behavior is well reproduced numerically on Figures 5 and 6. On Figure 7 one can verify that the difference between the experimental and simulated chromatograms decreases in amplitude as the discretization is refined.
We also display the solution of the measure equation computed with parameters of the isotherm slightly away from the best fit, for two discretizations $Nt = 1000$ on Figures 8 and 4000 on Figures 9. This last case corresponds to the discretization used to generate the "experimental" chromatogram, and the measure is a delta function when the target parameters are used (middle graphs b) for $N^* = 100$ and $K' = 5$). In fact derivatives computed for isotherms parameters away from the target value all look like delta functions but positioned at the wrong time. The Figure 10 displays a zoom of the derivatives obtained using the target parameters for the two discretizations. It corroborates clearly the convergence towards a delta function.

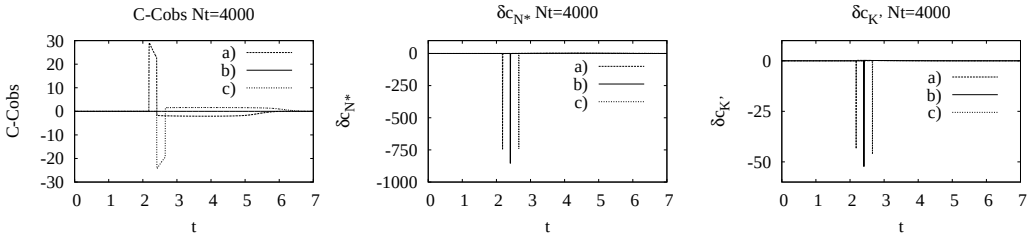

Figure 9: $c(L,t) - C_{obs}(t)$ and $\delta c(L,t)$ for $Nt = 4000$ time steps, and three parameter sets (N^*, K') a) $(90,4)$ b) $(100,5)$ and c) $(110,6)$.

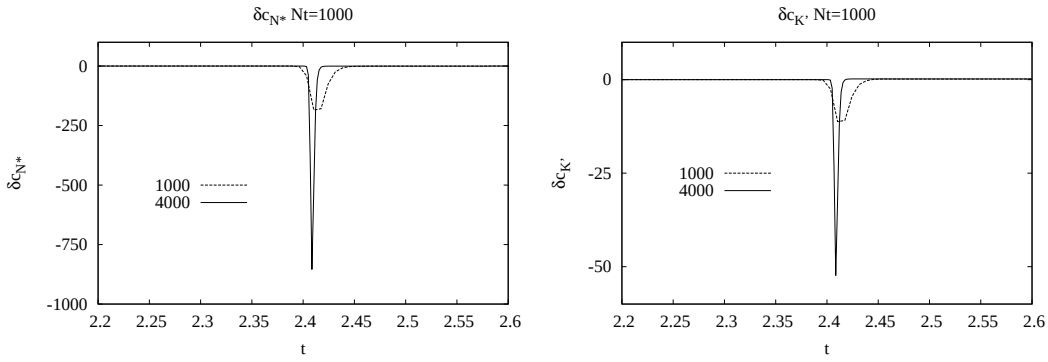

Figure 10: Zoom of $\delta c(L,t)$ for $Nt = 1000$ and 4000 time steps, and target parameters $(100, 5)$.

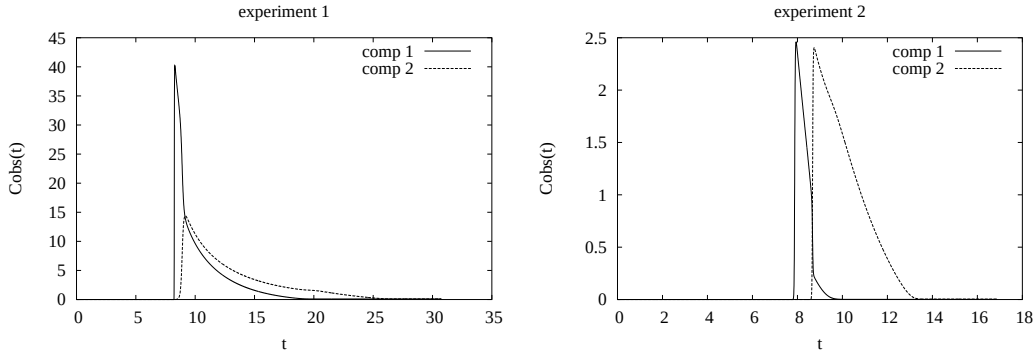

Figure 11: Chromatograms simulated with Table (3)

4.2 Binary mixture

We simulate in this case the two chromatograms displayed on figure 11 using a Bi-Langmuir isotherm for a two components mixture, namely

$$\mathbf{h}_i(\mathbf{c}) = N_1^* \frac{K_i^1 \mathbf{c}_i}{1 + K_1^1 \mathbf{c}_1 + K_2^1 \mathbf{c}_2} + N_2^* \frac{K_i^2 \mathbf{c}_i}{1 + K_1^2 \mathbf{c}_1 + K_2^2 \mathbf{c}_2}.$$

This model takes into account the fact that there are two possible kinds of adsorption sites, and is determined by 6 coefficients, whose values are listed in Table 3. We used a very refined discretization of 4000 time steps for each experiment, the second one until $T = 17$, with a flow rate of 1.2 ml/minutes the first and third one until $T = 31$, with a flow rate of 0.6 ml/minutes. The CFL condition is conservatively ensured by imposing

$$\frac{\Delta x}{\Delta t} \sup_{\mathbf{c}} \max_i |\lambda_i(\mathbf{c})| < 0.8,$$

where λ_i denote the eigenvalues of $\mathbf{F}'$ (see (6).

	component 1	component 2		
K_1	0.203564	0.283886	N_1^*	14.30
K_2	0.0325631	0.0407128	N_2^*	120.55

experiment	1		2		3	
component	1	2	1	2	1	2
injection	30.72	30.72	1.49	4.74	3.72	3.72

Table 3: Isotherm and injection parameters for the binary mixture simulation.

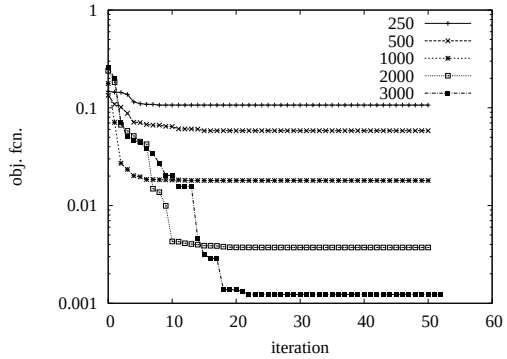

Figure 12: Convergence history for different discretizations, starting from $N^* = (10, 130)$.

We then perform identifications of the coefficients N_1^* and N_2^*, the coefficients K_1 and K_2 being kept equal to their theoretical values, which were used to generated the experimental data. Different discretizations and initial guesses are used. All the results tend to verify the robustness of the identification algorithm. The behavior of the objective function with the iterations of the minimization algorithm is displayed on Figure 12 for different discretizations. As expected the value of the objective function reached at convergence diminishes for finer discretization. Less predictably, this simulation shows that convergence is reached faster for coarser discretization. Figure 13 shows the distance of the parameters found by the minimization algorithm with the target parameters (used to generated the experimental chromatogram) There are two curves on this figure, corresponding to two different initial guesses $(N_1^*, N_2^*) = (130, 10)$ and $(110, 20)$. As expected, in both cases, the distance goes to zero when the discretization is refined. The parameters reached by the minimization algorithm are displayed for different discretizations and the two initial guesses on Figure 14. Even for coarse discretizations the dependence on the initial guess is very small.

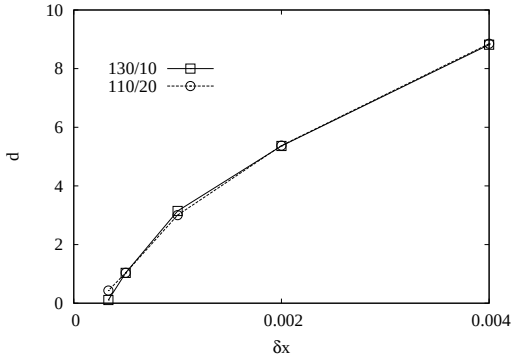

Figure 13: Distance to target at convergence as a function of the discretization.

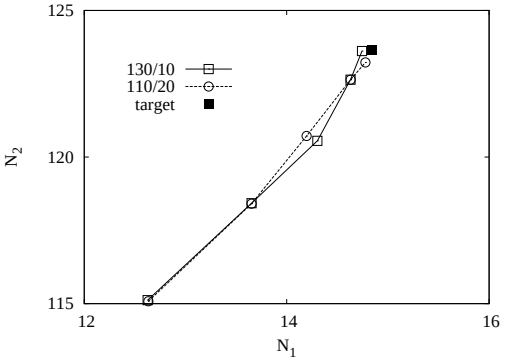

Figure 14: Target at convergence as a function of the discretization.

name / #	BA / 1	PE / 2	MBA / 3
K_i	0.01516	0.02341	0.02107
$K_i' = N_i^* K_i$	1.97	3.303	3.55
N_i^*	129.99	141.09	168.50

Table 4: Isotherm parameters for the ternary mixture [24].

5 Application on real datasets

5.1 Experimental identification of isotherms

We test the identification method on real datasets which were extensively studied by Quiñones, Ford and Guiochon in [24]. This paper is remarkable because it provides a very important set of isotherms measurements for a 3-components mixture: see Figure 15 below, where each symbol corresponds to a pair $\big(\mathbf{c}, \mathbf{h}(\mathbf{c})\big)$, and a different experimental setting. Let us explain briefly how to read this kind of figures. There are three components, 1 = BA = benzylalcohol, 2 = PE = 2−phenylethanol, 3 = MBA = 2−methylbenzylalcohol. For each component, $i = 1, 2, 3$, the adsorbed quantity $\mathbf{h}(\mathbf{c})_i$ is displayed as a function of the total amount of mixture $(c_1 + c_2 + c_3)$, for five different compositions of the mixture, namely $\diamond$ for single-component, $+$ for $3:1:1$ mixture[1], $\times$ for $1:1:1$, $\circ$ for $1:3:1$ and $\square$ for $1:1:3$ mixture.

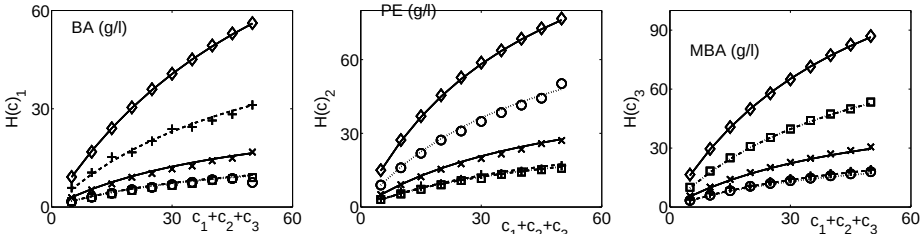

Figure 15: Isotherms experimental values and model values with the parameters of Table 4.

In this paper the authors make use of an isotherm function slightly different from the Langmuir isotherm, with six independent parameters

$$\mathbf{h}_i(\mathbf{c}) = N_i^* \frac{K_i \mathbf{c}_i}{1 + \sum_{i=1}^{p} K_i \mathbf{c}_i}. \tag{23}$$

They identify the coefficients K_i and N_i^* for $i = 1, \ldots, 3$ by fitting the curves obtained with (23) with the experimental points for single component mixtures ($\diamond$ in Figure 15), that is an amount of 30 experiments. The identified parameters are given in Table 4, and all the lines displayed in Figure 15 represent the isotherm computed with these values. Then simulated chromatograms are computed using (5), the modified isotherm

[1] 3 parts of BA, 1 part of PE and MBA

law (23) with the values obtained for the six parameters, and visually compared with a set of seven experimental chromatograms, corresponding to different proportions of the three components and different injection profiles. Two of them are displayed on Figures 16. The left-hand side corresponds to a mixture of BA, PE, and MPA with proportions $1:3:1$, while the right-hand side corresponds to a mixture of BA, PE, and MPA with proportions $3:1$. Experimental profiles are represented with symbols. The chromatograms computed with (5), and the values of parameters in Table 4 are represented with dashed lines ($--$), while solid lines (—) represent the chromatograms computed with the values in Table 5 below.
Notice that these computations are performed with an injection condition recorded from experimental data as well, see Figure 17, which is not a step function. The influence of this realistic boundary condition on the shape of the chromatograms is important, see [24] for a more detailed discussion. The left-hand injection profile corresponds to the $1:3:1$ experiment, the right-hand one to $3:1$.
Remark. From now on, all the numerical simulations we use are performed using the Godunov scheme, with 1000 space points, 1391 time steps, and the gradient is computed with the direct strategy described in Section 3.2.

5.2 Numerical identification of the isotherm

Now we bring our method into operation, using the two sets of chromatograms above as observation. This is a completely different identification strategy from the previous one, which is of course very reliable but unfortunately experimental estimates for isotherms are seldom available, while on the opposite chromatograms are more standard chromatography measurements.
In order to recover the coefficients with the same accuracy on each component and to use the information from both datasets we consider the following weighted cost function, where $\gamma_j = 1/\max_n \left|\mathbf{c}_{exp}(t_n)_j\right|$:

$$\tilde{J}(\alpha) = \frac{1}{2}\Delta t \sum_{\text{exp}=1,2} \sum_{j=1}^{3} \gamma_j \sum_{n=1}^{N} \left|\mathbf{c}_{K\,j}^{n} - \mathbf{c}_{exp}(t_n)_j\right|^2 . \tag{24}$$

As a first test, we start the descent with the parameters in Table 4 as an initial guess, the value of the cost function is then 0.115, and we optimize on the whole set of six parameters. After convergence of the conjugate gradient optimization, its value is 0.0107, it has been divided by ten and the parameters of the isotherm are displayed in Table 5. The fit with the experimental isotherm is displayed on Figure 15, the fit with experimental chromatograms on Figure 16 with solid lines (—), together with the simulations of [24] (dashed lines $--$). Several comments are in order here.
First, it is clear that the fit on the isotherms is worse for simulated parameters. This emphasize the fact that the kind of data we used to identify are completely different from [24]. The concentration range of the chromatograms is not the same as in the isotherm data, and in particular we have no information at all on the single-component adsorption ($\diamond$ in Figures 15 and 18), for which the simulated parameters are quite poor.

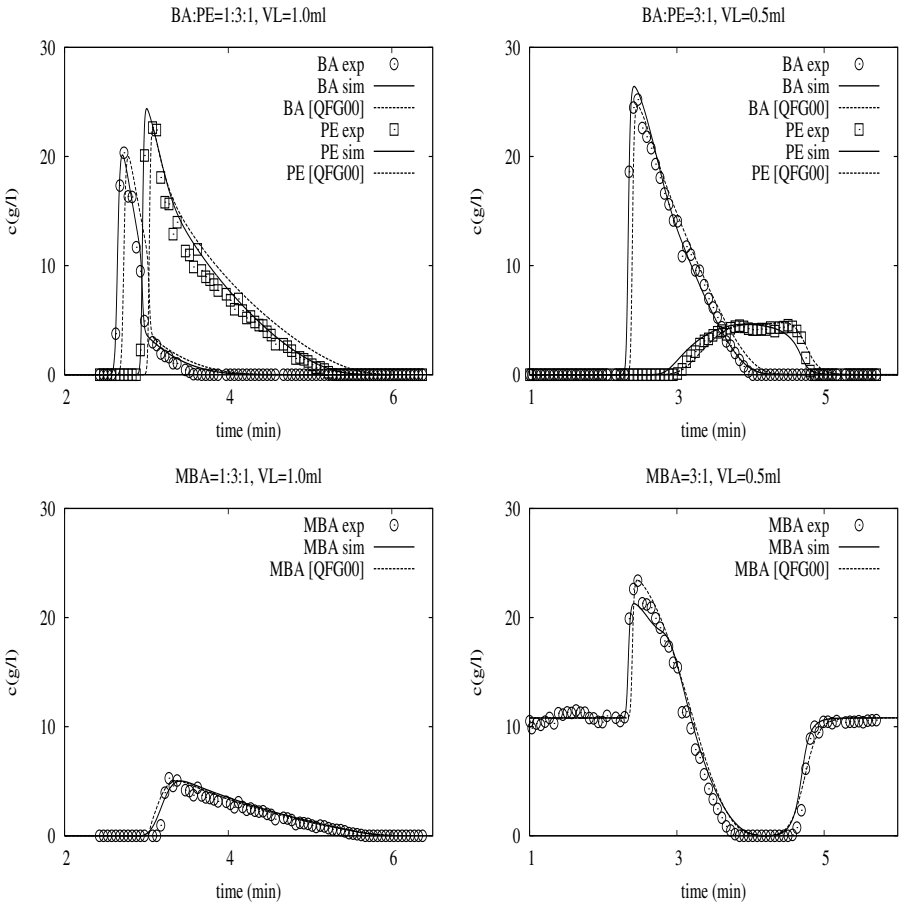

Figure 16: Experimental chromatograms (symbols) and simulated chromatograms using parameters from Table 4 (−−) and Table 5 (–). BA and PE components (top), MBA component (bottom).

name / #	BA / 1	PE / 2	MBA / 3
K_i	0.0137	0.0214	0.0206
$K_i' = N_i^* K_i$	1.78046	3.00974	3.47049
N_i^*	129.986	141.07	168.495

Table 5: Isotherm parameters after optimization starting from Table 4.

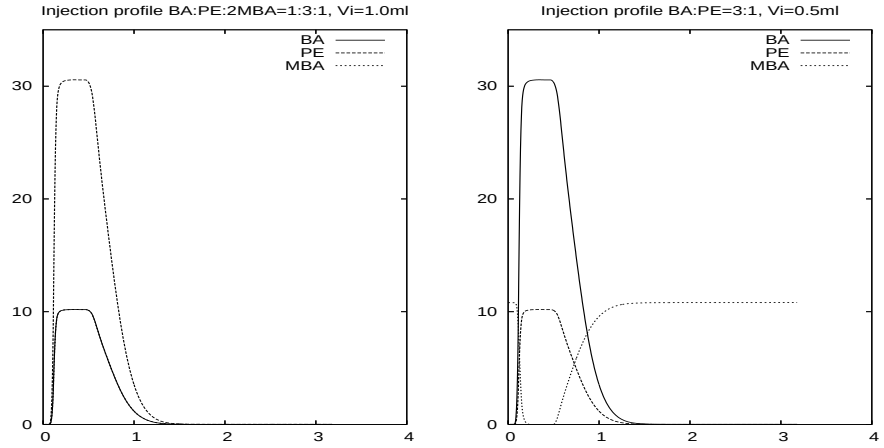

Figure 17: Injection profiles used for the two experiments.

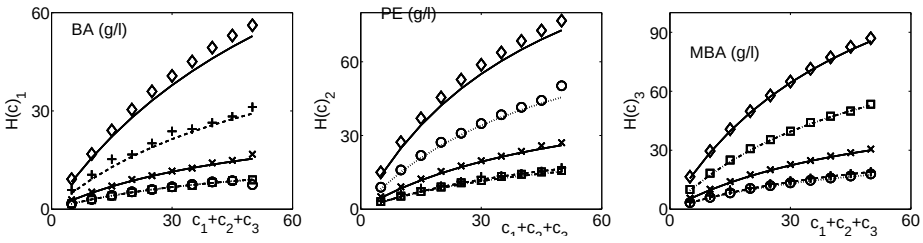

Figure 18: Isotherms experimental values and model values with the optimized parameters of Table 5.

name / #	BA / 1	PE / 2	MBA / 3
$K'_i = N^*_i K_i$	1.833	3.108	3.511

Table 6: Experimental values of K'_i, from [24]

starting point	$N^*_{BA/1}$	$N^*_{PE/2}$	$N^*_{MBA/3}$	initial cost value	final cost value
200. 200. 200.	120.937	134.14	169.401	0.206	0.00892
100. 100. 100.	123.724	135.768	158.81	0.411	0.00901
200. 200. 200.	124.059	135.537	159.29	0.432	0.00902
200. 100. 100.	123.342	135.968	157.67	0.292	0.00909
200. 100. 200.	*123.373*	*135.704*	*159.637*	*0.174*	*0.00896*
100. 200. 100.	120.003	133.503	173.244	0.122	0.00912
200. 200. 100.	124.167	135.949	157.592	0.237	0.00909
100. 100. 200.	118.536	133.099	178.772	0.210	0.00945
100. 150. 200.	116.171	133.053	182.999	0.046	0.00980
200. 150. 100.	124.46	136.054	156.623	0.071	0.00917
Mean value	**122.614**	**135.441**	**165.685**		
Relative variance	**0.09**	**0.030**	**0.45**		

Table 7: Isotherm parameters N^* identified from several starting points.

On the other hand, the value of the cost function has been divided by ten, and it is clear on Figure 16 that shocks are much better identified than from the experimental parameters. This is not very surprising, since all the numerical identification process is based upon the hyperbolic model, which is very sensitive to the shock position.
Next, we tried another series of identification, taking into account some experimental values for the parameters $K'_i = N^*_i K_i$, which are given in [24], and recalled in Table 6 (in the original paper, the displayed values are $K'_i \times (1-\epsilon)/\epsilon$, where $\epsilon = .59$ is the porosity of the medium). These values, corresponding to "analytical conditions", that is very small injected concentrations, are obtained with a good precision, and correspond to the propagation at concentration zero, which is given by $\mathbf{F}'(0)$. Therefore, we performed ten optimizations, keeping the values K'_i constant, and with different starting values for the three remaining parameters N^*_i. The results are gathered in Table 7. We notice that the cost function has been divided at least by a factor 10. The third coefficient seems to be more difficult to identify (relative variance 0.45), maybe because its concentration is rather small in one of the experiments. To give an illustration of this result, we display both the isotherm curves (Figure 19) and the chromatograms (Figure 20), for the fifth computation (the values are extracted from the italic line of Table 7, and recalled in Table 8). Concerning isotherms, the results are improved with respect to Figure 18, and this emphasize the fact that the experimental values of Table 6 include a significant information on isotherms in the computations. The results on chromatograms are good, in particular we still have the correct position of shocks, which was not so good in [24].

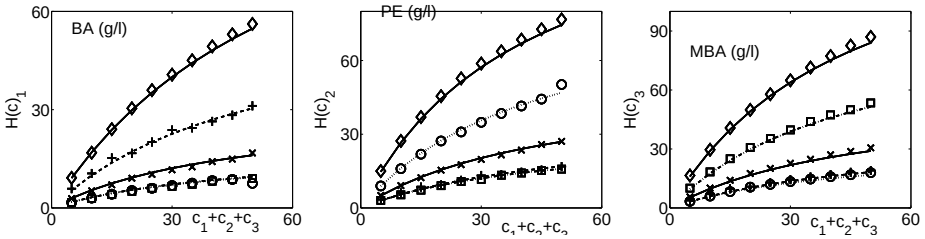

Figure 19: Isotherms experimental values and model values with the optimized parameters of Table 8.

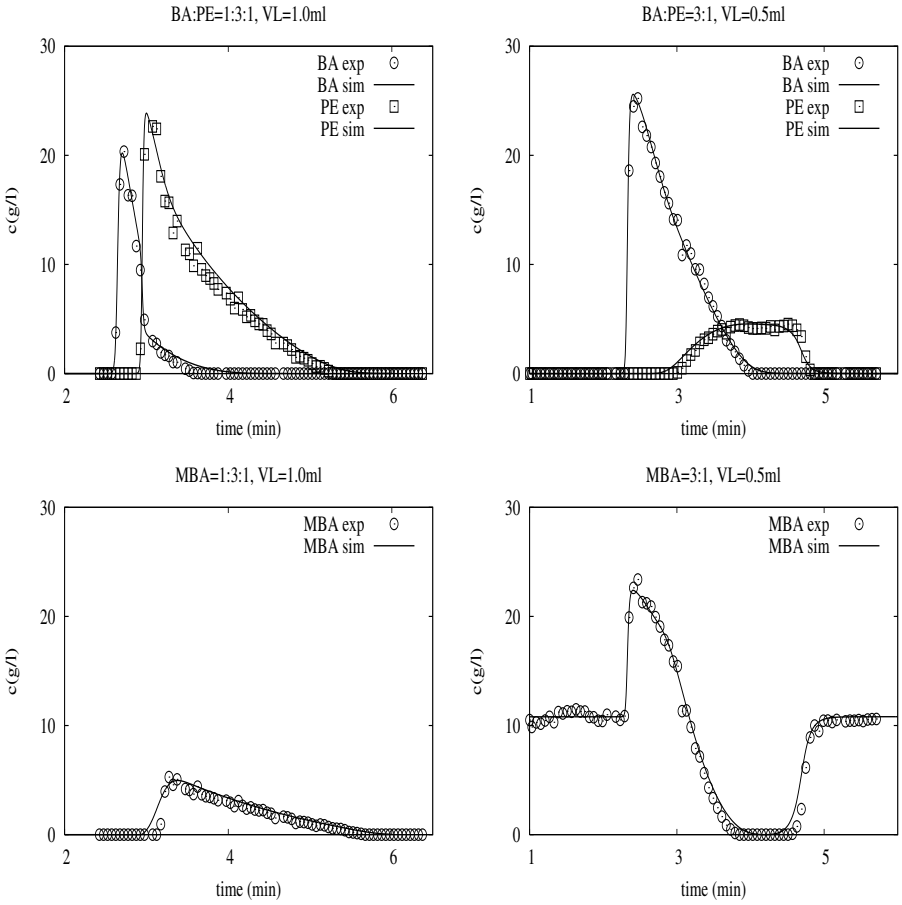

Figure 20: Experimental chromatograms and simulated chromatograms using optimized parameters from Table 8, BA and PE components (top) and MBA component (bottom).

name / #	BA / 1	PE / 2	MBA / 3
K_i	0.01486	0.0229	0.02199
$K_i' = N_i^* K_i$	1.833	3.108	3.511
N_i^*	123.373	135.704	159.637

Table 8: Isotherm parameters for the fifth experiment in Table 7.

6 Conclusion

The computation of the gradient required to use a descent optimization method in the flux identification for a system of conservation laws is now thoroughly studied from a numerical point of view. The direct computation of the gradient from the original problem, formulated with partial differential equations is still an open problem, even in the scalar case. However, convergence with respect to the discretization step is indicated by all the numerical tests. Moreover, the application of the identification method proved to be reliable in the case of a 3×3 system fully documented in the Chemical engineering literature. We emphasize that additional experimental information on the parameters can drastically improve the results.
On the other hand, it is clear that the complexity of the objective function makes it necessary to combine some global optimization method with the gradient, to avoid for instance local minima. A first idea is to use the global method to find a good starting point for the gradient (or any descent method), but more intricate couplings are currently under study, using evolutionary algorithms. Finally, the strategies for the formulation of the discrete gradient have to be applied to more complete models, taking into account for example diffusive effects, which are also widely used in Chemical Engineering.

Annex I. Adjoint formulation for the gradient estimation.

We describe in this annex the method to compute the gradient of the cost function (7) with a weak formulation. This method has been described and used in [25] and [16], but we give here the full computation, including the case of a non zero initial state, which is used in Section 5.
The gradient with respect to the model parameters will be derived through a Lagrangian formulation. We start with the development in the continuous case which gives a good feeling for the method. We will next present the discrete Lagrangian model which is actually used in the numerical simulations.

Gradient computation in the continuous case

The constraint for our optimization problem is that

$$\mathbf{c}_\alpha \text{ is solution to (5).}$$

In order to write a Lagrangian for this constraint, we first write a weak formulation:

$E(\mathbf{c},\mathbf{p},\alpha)=0, \quad \forall \mathbf{p}$, where the functional E is obtained by multiplying the PDE in (5) by a test function $\mathbf{p}$, as smooth as $\mathbf{c}$, and integrating by parts on the domain $[0,L]\times[0,T]$:

$$\begin{aligned}
E(\mathbf{c},\mathbf{p},\alpha) &= \int_0^L\int_0^T < \left(\partial_x\mathbf{c}+\partial_t\mathbf{F}(\mathbf{c})\right),\mathbf{p} > dt\,dx \\
&= \int_0^T\left([< \mathbf{c},\mathbf{p} >]_0^L - \int_0^L < \mathbf{c},\partial_x\mathbf{p}dx >\right)dt \\
&\quad + \int_0^L\left([< \mathbf{F}(\mathbf{c}),\mathbf{p} >]_0^T - \int_0^T < \mathbf{F}(\mathbf{c}),\partial_t\mathbf{p} > dt\right)dx \\
&= \int_0^T < \mathbf{c}(L,t),\mathbf{p}(L,t) > dt - \int_0^T < \mathbf{c}_{inj}(t),\mathbf{p}(0,t) > dt \\
&\quad + \int_0^L < \mathbf{F}(\mathbf{c}(x,T)),\mathbf{p}(x,T) > dx - \int_0^L < \mathbf{F}(\mathbf{c}_{ini}(x)),\mathbf{p}(x,0) > dx \\
&\quad - \int_0^L\int_0^T (< \mathbf{c},\partial_x\mathbf{p} > + < \mathbf{F}(\mathbf{c}),\partial_t\mathbf{p} >)\,dt\,dx.
\end{aligned}$$

The Lagrangian for the constrained minimization problem is

$$L(\mathbf{c},\mathbf{p},\alpha) = J(\mathbf{c}) - E(\mathbf{c},\mathbf{p},\alpha). \tag{25}$$

We notice that

$$\tilde{J}(\alpha) = J(\mathbf{c}_\alpha) = L(\mathbf{c}_\alpha,\mathbf{p},\alpha),$$

so that when we apply formally the chain rule, we obtain

$$J'(\alpha)\delta\alpha = \frac{\partial L}{\partial \mathbf{c}}(\mathbf{c}_\alpha,\mathbf{p},\alpha)\frac{\partial \mathbf{c}}{\partial \mathbf{F}}\,\delta\alpha + \frac{\partial L}{\partial \alpha}(\mathbf{c}_\alpha,\mathbf{p},\alpha)\,\delta\alpha.$$

Since it is hard to compute $\frac{\partial \mathbf{c}}{\partial \alpha}\delta\alpha$ (for reasons mentioned in paragraph 3) we will choose $\mathbf{p}$ such that $\frac{\partial L}{\partial \mathbf{c}}(\mathbf{c}_\alpha,\mathbf{p},\alpha)=0$ and will next compute $\tilde{J}'(\alpha)\delta\alpha = \frac{\partial L}{\partial \alpha}(\mathbf{c}_\alpha,\mathbf{p},\alpha)\,\delta\alpha$ for this special $\mathbf{p}$.
When we differentiate the Lagrangian with respect to $\mathbf{c}$, the terms involving $\mathbf{c}_{inj}$ and $\mathbf{c}_{ini}$ disappear, because they are fixed data of the problem, the term $< \mathbf{F}(\mathbf{c}),\partial_t\mathbf{p} >$ leads to $\langle \partial_{\mathbf{c}}\mathbf{F}(\mathbf{c})\delta\mathbf{c},\partial_t\mathbf{p}\rangle = \langle \delta\mathbf{c},\left(\partial_{\mathbf{c}}F(\mathbf{c})\right)^T\partial_t\mathbf{p}\rangle$, so that, for the cost function (7), we have

$$\begin{aligned}
\frac{\partial L}{\partial \mathbf{c}} &= \frac{\partial J}{\partial \mathbf{c}} - \frac{\partial E}{\partial \mathbf{c}} \\
&= \int_0^T \langle(\mathbf{c}(L,t)-\mathbf{c}_{obs}(t)),\delta\mathbf{c}\rangle dt - \int_0^T \langle \delta\mathbf{c},\mathbf{p}(L,t)\rangle dt \\
&\quad - \int_0^L \langle \delta\mathbf{c},\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}(x,T))^T\mathbf{p}(x,T)\rangle\,dx \\
&\quad + \int_0^L\int_0^T \langle \delta\mathbf{c},\left(\partial_x\mathbf{p}+\partial_{\mathbf{c}}\mathbf{p}(\mathbf{c})^T\partial_t\mathbf{p}\right)\rangle\,dt\,dx.
\end{aligned}$$

Putting this equal to 0 for all possible $\delta\mathbf{c}$ can be interpreted as a weak formulation of a linear transport equation for $\mathbf{p}$, with a boundary condition on $x = L$, readily given by the first two terms in this formula, and a final datum on $t = T$, which reads $\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}(x,T))^T\mathbf{p}(x,T) = 0$. But since $\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}(x,T))$ is invertible, we can finally choose the adjoint $\mathbf{p}$ solution to (18).
A formula for the gradient of $\tilde{J}(\alpha)$ is then obtained by picking any $\mathbf{p}$ solution to (18) and computing

$$\begin{aligned}\tilde{J}(\alpha)\delta\alpha &= \frac{\partial L}{\partial\alpha}\delta\alpha = -\frac{\partial E}{\partial\alpha}\delta\alpha \\ &= \int_0^L \delta\mathbf{F}(\mathbf{c}_{ini}(x))\mathbf{p}(x,0)\,dx - \int_0^L \delta\mathbf{F}(\mathbf{c}(x,T))\mathbf{p}(x,T)\,dx \\ &\quad + \int_0^L\int_0^T \delta\mathbf{F}(\mathbf{c})\partial_t\mathbf{p}\,dt\,dx.\end{aligned}$$

$\delta\mathbf{F}$ is obtained by differentiating $\mathbf{F}$ with respect to the parameters we want to identify (see (16) and annex III).

Discrete formulation

We first write the discrete analogue of formula (25) for the numerical scheme (8)

$$\begin{aligned}\tilde{E}(\mathbf{c},\mathbf{a},\alpha) &= \sum_{n=1}^{N}\sum_{k=0}^{K-1}\langle \mathbf{c}_{k+1}^n - \mathbf{c}_k^n + \lambda\left(\mathbf{F}(\mathbf{c}_k^n) - \mathbf{F}(\mathbf{c}_k^{n-1})\right), \mathbf{a}_{k+1}^{n-1}\rangle \\ &= \sum_{n=1}^{N}\sum_{k=0}^{K-1}\langle \mathbf{c}_{k+1}^n, \mathbf{a}_{k+1}^{n-1}\rangle - \sum_{n=1}^{N}\sum_{k=0}^{K-1}\langle \mathbf{c}_k^n, \mathbf{a}_{k+1}^{n-1}\rangle \\ &\quad + \lambda\sum_{n=1}^{N}\sum_{k=0}^{K-1}\langle \mathbf{F}(\mathbf{c}_k^n), \mathbf{a}_{k+1}^{n-1}\rangle - \lambda\sum_{n=1}^{N}\sum_{k=0}^{K-1}\langle \mathbf{F}(\mathbf{c}_k^{n-1}), \mathbf{a}_{k+1}^{n-1}\rangle. \qquad (26)\end{aligned}$$

Since we will not differentiate with respect to the terms $\mathbf{c}_0^n$ which are fixed by the injection condition, we set them apart. Furthermore, we move up the k indexes in the first sum and the n indexes in the fourth sum:

$$\begin{aligned}\tilde{E}(\mathbf{c},\mathbf{a},\alpha) &= \sum_{n=1}^{N}\sum_{k=1}^{K}\langle \mathbf{c}_k^n, \mathbf{a}_k^{n-1}\rangle - \sum_{n=1}^{N}\sum_{k=0}^{K-1}\langle \mathbf{c}_k^n, \mathbf{a}_{k+1}^{n-1}\rangle \\ &\quad + \lambda\sum_{n=1}^{N}\sum_{k=0}^{K-1}\langle \mathbf{F}(\mathbf{c}_k^n), \mathbf{a}_{k+1}^{n-1}\rangle - \lambda\sum_{n=0}^{N-1}\sum_{k=0}^{K-1}\langle \mathbf{F}(\mathbf{c}_k^n), \mathbf{a}_{k+1}^{n}\rangle \\ &= \sum_{n=1}^{N-1}\sum_{k=1}^{K-1}\left\{\langle \mathbf{c}_k^n, \mathbf{a}_k^{n-1}\rangle - \langle \mathbf{c}_k^n, \mathbf{a}_{k+1}^{n-1}\rangle + \lambda\left(\langle \mathbf{F}(\mathbf{c}_k^n), \mathbf{a}_{k+1}^{n-1}\rangle - \langle \mathbf{F}(\mathbf{c}_k^n), \mathbf{a}_{k+1}^{n}\rangle\right)\right\} \\ &\quad + \sum_{k=1}^{K}\langle \mathbf{c}_k^N, \mathbf{a}_k^{N-1}\rangle + \sum_{n=1}^{N-1}\langle \mathbf{c}_K^n, \mathbf{a}_K^{n-1}\rangle - \sum_{k=0}^{K-1}\langle \mathbf{c}_k^N, \mathbf{a}_{k+1}^{N-1}\rangle - \sum_{n=1}^{N-1}\langle \mathbf{c}_0^n, \mathbf{a}_1^{n-1}\rangle\end{aligned}$$

$$+ \quad \lambda\left(\sum_{k=0}^{K-1}\langle \mathbf{F}(\mathbf{c}_k^N), \mathbf{a}_{k+1}^{N-1}\rangle + \sum_{n=1}^{N-1}\langle \mathbf{F}(\mathbf{c}_0^n), \mathbf{a}_1^{n-1}\rangle - \sum_{k=0}^{K-1}\langle \mathbf{F}(\mathbf{c}_k^0), \mathbf{a}_{k+1}^0\rangle - \sum_{n=1}^{N-1}\langle \mathbf{F}(\mathbf{c}_0^n), \mathbf{a}_1^n\rangle\right).$$

As in the continuous case, we write the discrete Lagrangian where the constraint

$$\left(\mathbf{c}_k^n\right) \text{ solution of (8)}$$

is taken into account by $\tilde{E}(\mathbf{c}, \mathbf{a}, \alpha) = 0$. We are led to differentiate $\tilde{L}(\mathbf{c}, \mathbf{a}, \alpha) = \tilde{J}(\mathbf{c}) - \tilde{E}(\mathbf{c}, \mathbf{a}, \alpha)$ with respect to all components of $\left(\mathbf{c}_k^n\right)$. One the one hand, we get for the cost function

$$\frac{\partial \tilde{J}}{\partial \mathbf{c}_k^n} = \begin{cases} \Delta t\langle\left(\mathbf{c}_K^n - \mathbf{c}_{exp}(t_n)\right), \delta\mathbf{c}_k^n\rangle, & 0 \le n \le N, \\ 0, & 0 \le k < K, \quad 0 \le n \le N. \end{cases}$$

On the other hand, the derivatives of the weak formulation are given by

$$\begin{aligned} \frac{\partial \tilde{E}}{\partial \mathbf{c}_k^n} &= \langle \delta\mathbf{c}_k^n, \mathbf{a}_k^{n-1} - \mathbf{a}_{k+1}^{n-1}\rangle + \lambda\langle \partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^n)\delta\mathbf{c}_k^n, \left(\mathbf{a}_{k+1}^{n-1} - \mathbf{a}_{k+1}^n\right)\rangle \\ &= \langle \delta\mathbf{c}_k^n, \mathbf{a}_k^{n-1} - \mathbf{a}_{k+1}^{n-1} + \lambda\left(\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^n)\right)^T\left(\mathbf{a}_{k+1}^{n-1} - \mathbf{a}_{k+1}^n\right)\rangle \end{aligned}$$

for $k = 1, \ldots, K-1$ and $n = 1, \ldots, N-1$,

$$\frac{\partial \tilde{E}}{\partial \mathbf{c}_K^n} = \langle \delta\mathbf{c}_K^n, \mathbf{a}_K^{n-1}\rangle, \qquad 1 \le n \le N-1,$$

$$\frac{\partial \tilde{E}}{\partial \mathbf{c}_k^N} = \langle \delta\mathbf{c}_k^N, \mathbf{a}_k^{N-1} - \mathbf{a}_{k+1}^{N-1} + \lambda\left(\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^N)\right)^T\mathbf{a}_{k+1}^{N-1}\rangle, \qquad 1 \le k \le K-1.$$

Imposing that all partial derivatives of $\tilde{L} = \tilde{J} - \tilde{E}$ with respect to $\mathbf{c}_k^n$ must be zero, we obtain the following formulæ for $\mathbf{a}_k^n$:

$$\begin{cases} \mathbf{a}_k^{n-1} = \mathbf{a}_{k+1}^{n-1} - \lambda\left(\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^n)\right)^T\left(\mathbf{a}_{k+1}^{n-1} - \mathbf{a}_{k+1}^n\right), & 0 \le k < K, \quad 0 < n \le N-1, \\ \mathbf{a}_K^n = \Delta t\left(\mathbf{c}_K^{n+1} - \mathbf{c}_{exp}(t_{n+1})\right), & 0 < n < N, \\ \mathbf{a}_k^{N-1} = \mathbf{a}_{k+1}^{N-1} - \lambda\left(\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^N)\right)^T\mathbf{a}_{k+1}^{N-1}, & 1 \le k \le K-1. \end{cases} \tag{27}$$

In order to recover a discretization scheme for the continuous backward equation (18), compatible with the final condition at $t = T$, we are led to impose $\mathbf{a}_k^N = 0$ for all $0 < k \le K$, so that the third relation in (27) rewrites exactly as the first one, for $0 < n \le N$ and $0 \le k < K$. Thus we obtain (22).

The gradient of $\tilde{J}$ for this adjoint is next computed by plugging it into (26) and differentiating $\tilde{E}$ with respect to α. Only the two terms depending on $\mathbf{F}$ plays a role. We obtain

$$\begin{aligned} \tilde{J}'(\alpha)\delta\alpha &= -\frac{\partial E}{\partial \alpha}\delta\alpha \\ &= -\lambda\sum_{n=1}^{N}\sum_{k=0}^{K-1}\delta\mathbf{F}(\mathbf{c}_k^n)\left(\mathbf{a}_{k+1}^{n-1} - \mathbf{a}_{k+1}^n\right) + \lambda\sum_{k=0}^{K-1}\delta\mathbf{F}(\mathbf{c}_k^0)\mathbf{a}_{k+1}^0, \end{aligned}$$

where $\delta\mathbf{F} = \partial_\alpha\mathbf{F}\delta\alpha$, (see (16) and Annex III below).

Annex II. Proof of Lemma 1

We study the limit of the Newton ratio

$$\frac{\tilde{J}(\mathbf{F}+\alpha\delta\mathbf{F})-\tilde{J}(\mathbf{F})}{\alpha}=\Delta t\sum_{n=0}^{N}\left\langle\frac{\mathbf{d}_K^n-\mathbf{c}_K^n}{\alpha},\frac{\mathbf{d}_K^n+\mathbf{c}_K^n}{2}-\mathbf{c}_{exp}(t_n)\right\rangle.$$

where $\mathbf{d}_k^n$ is the solution of the Godunov scheme (8) associated to the perturbed flux $\mathbf{F}+\alpha\delta\mathbf{F}$

$$\mathbf{d}_{k+1}^n=\mathbf{d}_k^n-\lambda\left((\mathbf{F}+\alpha\delta\mathbf{F})(\mathbf{d}_k^n)-(\mathbf{F}+\alpha\delta\mathbf{F})(\mathbf{d}_k^{n-1})\right).$$

We actually prove that uniformly in $k=0,\dots,K$

$$\lim_{\alpha\to 0}\|\mathbf{d}_k-\mathbf{c}_k\|=0\qquad\text{and}\quad\lim_{\alpha\to 0}\|\frac{\mathbf{d}_k-\mathbf{c}_k}{\alpha}-\delta\mathbf{c}_k^n\|=0,\tag{28}$$

where $\delta\mathbf{c}_k^n$ is given by the scheme (21) and

$$\|\mathbf{c}_k\|=\max_{n=0,\dots,N}\left|\mathbf{c}_k^n\right|.$$

We set $\mathbf{r}_k^n=\mathbf{d}_k^n-\mathbf{c}_k^n$, which verifies

$$\begin{cases}\mathbf{r}_{k+1}^n-\mathbf{r}_k^n+\lambda\left((\mathbf{F}+\alpha\delta\mathbf{F})(\mathbf{d}_k^n)-(\mathbf{F}+\alpha\delta\mathbf{F})(\mathbf{d}_k^{n-1})\right)-\lambda\left(\mathbf{F}(\mathbf{c}_k^n)-\mathbf{F}(\mathbf{c}_k^{n-1})\right)=0,\\ \mathbf{r}_0^n=0,\\ \mathbf{r}_k^0=0.\end{cases}$$

After some algebra, and provided that $\mathbf{F}$ is smooth enough, we obtain that the quantity $\mathbf{z}_k^n=\mathbf{r}_k^n/\alpha$ satisfies

$$\begin{cases}\mathbf{z}_{k+1}^n-\left(1-\lambda\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^n)\right)\mathbf{z}_k^n-\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^{n-1})\mathbf{z}_k^{n-1}+\lambda\left(\delta\mathbf{F}(\mathbf{d}_k^n)-\delta\mathbf{F}(\mathbf{d}_k^{n-1})\right)=O(\alpha,\|\mathbf{z}_k\|^2),\\ \mathbf{z}_0^n=0,\\ \mathbf{z}_k^0=0.\end{cases}\tag{29}$$

We prove now that $\|\mathbf{z}^n\|\le M$. Provided that λ verifies a CFL condition $\lambda<1/\|\partial_{\mathbf{c}}\mathbf{F}\|$, we have from (29)

$$\|\mathbf{z}_{k+1}\|\le\|\mathbf{z}_k\|+a\|\mathbf{z}_k\|^2+b,$$

with $a=C\lambda\alpha$ where C is a bound of $\|\partial_{\mathbf{cc}}\mathbf{F}\|$ and $b=2\lambda\|\delta\mathbf{F}\|$. Using $\|\mathbf{z}_0\|=0$ we obtain by summation

$$\|\mathbf{z}_K\|\le a\sum_{k=0}^{K-1}\|\mathbf{z}_k\|^2+Kb.$$

In order to bound $\|\mathbf{z}_k\|$ by M we therefore need to have

$$aKM^2-M+Kb\le 0\tag{30}$$

which can be true if $4abK^2 < 1$, in other words if

$$\alpha < \frac{1}{8\lambda^2 C\|\delta\mathbf{F}\|K^2}.$$

In that case the smallest root of (30) provides a bound for $\|\mathbf{z}_k\|$

$$M = \frac{1 - \sqrt{1 - 4abK^2}}{2aK} = \lambda\|\delta\mathbf{F}\|K + O(\alpha) < \frac{\lambda}{2}\|\delta\mathbf{F}\|K,$$

here again, provided that α is sufficiently small. Since $\mathbf{r}_k^n = \alpha\mathbf{z}_k^n$, this proves the first limit in (28).
Now, for the difference $\Delta_k^n = \mathbf{z}_k^n - \delta\mathbf{c}_k^n$, we obtain in the same way, using the bound on $\|\mathbf{z}_k\|$, and if $\delta\mathbf{F}$ is smooth enough,

$$\Delta_{k+1}^n - \Delta_k^n\left(1 - \lambda\partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^n)\right) \quad - \quad \partial_{\mathbf{c}}\mathbf{F}(\mathbf{c}_k^{n-1})\Delta_k^{n-1} = O(\alpha).$$

We use again $\lambda < 1/\|\partial_{\mathbf{c}}\mathbf{F}\|$ to obtain that, for α small enough,

$$\|\Delta_{k+1}\| \leq \|\Delta_k\| + O(\alpha).$$

Since $\|\Delta_0\| = 0$, this gives in turn

$$\|\Delta_k\| \leq O(\alpha).$$

Therefore, for a given discretization, $\|\Delta_k\| \to 0$ when $\alpha \to 0$. □

Annex III. Gradient of isotherms

We have put together in this annex the explicit formulæ to compute the $\delta\mathbf{F}$ involved in (28) or (20–21), for several isotherms functions $\mathbf{h}$ (see (16) and (6)).
In the case of Langmuir isotherm (3), one can optimize with respect to the parameters N^* and $(K_i)_{i=1}^p$. In that case the partial derivatives are

$$\left(\frac{\partial\mathbf{h}(\mathbf{c})}{\partial N^*}\right)_i = \frac{K_i\mathbf{c}_i}{1 + \sum_{j=1}^p K_j\mathbf{c}_j}, \quad \left(\frac{\partial\mathbf{h}(\mathbf{c})}{\partial K_l}\right)_i = N^*\frac{\delta_{li}\mathbf{c}_i}{1 + \sum_{j=1}^p K_j\mathbf{c}_i} - N^*\frac{K_i\mathbf{c}_i\mathbf{c}_l}{(1 + \sum_{j=1}^p K_j\mathbf{c}_j)^2}. \tag{31}$$

One can also write this isotherm as a function of N^* and $(K_i' = N^*K_i)_{i=1}^p$.

$$\mathbf{h}_i(\mathbf{c}) = N^*\frac{K_i'\mathbf{c}_i}{N^* + \sum_{j=1}^p K_j'\mathbf{c}_j}.$$

This last option is often more interesting because experimental chromatograms provide us with a reliable and direct estimation of the K_i' values. Partial derivatives with respect

to N^* and K'_i are

$$\left(\frac{\partial \mathbf{h}(\mathbf{c})}{\partial N^*}\right)_i = \frac{K'_i\mathbf{c}_i \sum_{j=1}^{p} K'_j\mathbf{c}_j}{(N^* + \sum_{j=1}^{p} K'_j\mathbf{c}_j)^2}, \quad \left(\frac{\partial \mathbf{h}(\mathbf{c})}{\partial K'_l}\right)_i = N^* \frac{\delta_{li}\mathbf{c}_i}{N^* + \sum_{j=1}^{p} K'_j\mathbf{c}_i} - N^* \frac{K'_i\mathbf{c}_i\mathbf{c}_l}{(N^* + \sum_{j=1}^{p} K'_j\mathbf{c}_j)^2}. \quad (32)$$

The Bi-Langmuir isotherm (23) used to validate the binary mixture identification has the following partial derivatives

$$\left(\frac{\partial \mathbf{h}(\mathbf{c})}{\partial N^*_l}\right)_i = \frac{K^l_i\mathbf{c}_i}{1 + \sum_{j=1}^{p} K^l_j\mathbf{c}_j}, \quad \left(\frac{\partial \mathbf{h}(\mathbf{c})}{\partial K^m_l}\right)_i = N^*_m \frac{\delta_{li}\mathbf{c}_i}{1 + \sum_{j=1}^{p} K^m_j\mathbf{c}_i} - N^*_m \frac{K^m_i\mathbf{c}_i\mathbf{c}_l}{(1 + \sum_{j=1}^{p} K^m_j\mathbf{c}_j)^2}. \quad (33)$$

In the case of the isotherm (23) used to model the experimental data there are $2p$ parameters $(N^*_i)_{i=1,\dots,p}$ and $(K_i)_{i=1,\dots,p}$ versus which the derivatives are the following

$$\left(\frac{\partial \mathbf{h}(\mathbf{c})}{\partial N^*_l}\right)_i = \delta_{li} \frac{K_i\mathbf{c}_i}{1 + \sum_{j=1}^{p} K_j\mathbf{c}_j}, \quad \left(\frac{\partial \mathbf{h}(\mathbf{c})}{\partial K_l}\right)_i = N^*_i \frac{\delta_{li}\mathbf{c}_i}{1 + \sum_{j=1}^{p} K_j\mathbf{c}_i} - N^*_i \frac{K_i\mathbf{c}_i\mathbf{c}_l}{(1 + \sum_{j=1}^{p} K_j\mathbf{c}_j)^2}. \quad (34)$$

As done in the Langmuir case, this isotherm can be rewritten as a function of the $(N^*_i)_{i=1}^p$ and $(K'_i = N^*_i K_i)_{i=1}^p$

$$\mathbf{h}_i(\mathbf{c}) = \frac{K'_i\mathbf{c}_i}{1 + \sum_{j=1}^{p} \frac{K'_j\mathbf{c}_j}{N^*_j}}.$$

Partial derivatives with respect to N^*_i and K'_i are

$$\left(\frac{\partial \mathbf{h}(\mathbf{c})}{\partial N^*_l}\right)_i = \frac{K'_iK'_l\mathbf{c}_i\mathbf{c}_l}{N^{*2}_l\left(1 + \sum_{j=1}^{p} \frac{K'_j\mathbf{c}_j}{N^*_j}\right)^2}, \quad \left(\frac{\partial \mathbf{h}(\mathbf{c})}{\partial K'_l}\right)_i = \frac{\delta_{li}\mathbf{c}_i}{1 + \sum_{j=1}^{p} \frac{K'_j\mathbf{c}_j}{N^*_j}} - \frac{K'_i\mathbf{c}_i\mathbf{c}_l}{N^*_l\left(1 + \sum_{j=1}^{p} \frac{K'_j\mathbf{c}_j}{N^*_j}\right)^2}. \quad (35)$$

References

[1] L. Ambrosio. Transport equation and cauchy problem for bv vector fields. *Invent. Math.*, 158(2):227–260, 2004.

[2] C. Bardos and O. Pironneau. A formalism for the differentiation of conservation laws. *C. R. Math. Acad. Sci. Paris*, 335(10):839–845, 2002.

[3] Ch. Bernardi and O. Pironneau. Derivative with respect to discontinuities in the porosity. *C. R. Math. Acad. Sci. Paris*, 335(7):661–666, 2002.

[4] F. Bouchut and F. James. One-dimensional transport equations with discontinuous coefficients. *Nonlinear Analysis, TMA*, 32(7):891–933, 1998.

[5] F. Bouchut and F. James. Differentiability with respect to the initial data for a scalar conservation law. In *Hyperbolic problems: theory, numerics, applications, Vol. I (Zürich, 1998)*, volume 129 of *Internat. Ser. Numer. Math.*, pages 113–118. Birkhäuser, Basel, 1999.

[6] F. Bouchut, F. James, and S. Mancini. Uniqueness and weak stability for multidimensional transport equations with one-sided lipschitz coefficients. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5)*, IV:1–25, 2005.

[7] F. Bouchut and B. Perthame. Kružkov's estimates for scalar conservation laws revisited. *Trans. Amer. Math. Soc.*, 350(7):2847–2870, 1998.

[8] R.J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98:511–547, 1989.

[9] A. E. Eiben and M. Schoenauer, editors. *Special issue on evolutionary computing*. Elsevier Science Publishers B.V., Amsterdam, 2002. Inform. Process. Lett. **82** (2002), no. 1.

[10] S. Evje, K. H. Karlsen, and N. H. Risebro. A continuous dependence result for nonlinear degenerate parabolic equations with spatially dependent flux function. In *In H. Freistühler and G. Warnecke, editors, "Hyperbolic problems: theory, numerics, applications"*, volume 140 of *Internat. Ser. Numer. Math.*, pages 337–346. Birkhäuser, 2001.

[11] A. Fadda and M. Schoenauer. Evolutionary chromatographic law identification by recurrent neural nets. In EP95editors, editor, *EP95*, pages 219–235. MIT Press, March 1995.

[12] E. Godlewski and P.A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws*, volume 118 of *Applied Mathematical Sciences*. Springer Verlag, New-York, USA, 1996.

[13] L. Gosse and F. James. Numerical approximations of one-dimensional linear conservation equations with discontinuous coefficients. *Math. Comp.*, 69:987–1015, 2000.

[14] G. Guiochon, A. Feilinger, S. Golshan Shirazi, and A. Katti. *Fundamentals of preparative and nonlinear chromatography*. Academic Press, Boston, second edition, 2006.

[15] G. Guiochon, S. Golshan Shirazi, and A. Katti. *Fundamentals of preparative and nonlinear chromatography*. Academic Press, Boston, 1994.

[16] F. James and M. Sepúlveda. Parameter identification for a model of chromatographic column. *Inverse Problems*, 10(6):1299–1314, 1994.

[17] F. James and M. Sepúlveda. Convergence results for the flux identification in a scalar conservation law. *SIAM J. Control & Opt.*, 37(3):869–891, 1999.

[18] F. James, M. Sepúlveda, I. Quiñones, F. Charton, and G. Guiochon. Determination of binary competitive equilibrium isotherms from the individual chromatographic band profiles. *Chem. Eng. Sci.*, 54(11):1677–1696, 1999.

[19] I. Langmuir. The constitution and fundamental properties of solids and liquids. part i. solids. *J. Amer. Chem. Soc.*, 38:2221–2295, 1916.

[20] I. Langmuir. The adsorption of gases on plane surfaces of glass, mica and platinum. *J. Amer. Chem. Soc.*, 40(9):1361–1403, 1918.

[21] B. Lucier. A moving mesh numerical method for hyperbolic conservation laws. *Math. of Comp.*, 173:59–69, 1986.

[22] O. Pironneau and E. Polak. Consistent approximations and approximate functions and gradients in optimal control. *SIAM J. Control Optim.*, 41(2):487–510 (electronic), 2002.

[23] F. Poupaud and M. Rascle. Measure solutions to the linear multidimensional transport equation with discontinuous coefficients. *Comm. Partial Diff. Equ.*, 22:337–358, 1997.

[24] I. Quiñones, J. C. Ford, and G. Guiochon. High concentration band profiles and system peaks for a ternary solute system. *Anal. Chem.*, 72:1495–1502, 2000.

[25] M. Sepúlveda. Identification de paramètres pour un système hyperbolique. application à l'estimation des isothermes en chromatographie. Thèse de doctorat, Ecole Polytechnique, France, 1993.